

Academic suicide in South Africa?

The outspoken universities of South Africa are under new threat from the government that supports them. Academics elsewhere must rally round, in their own interests if not South Africa's.

THE government of South Africa has moved with its customary guile against the universities it supports. Earlier in the year, the government decreed that universities must impose sanctions on students and teachers who engage in public demonstrations against the government, and against its policy of apartheid in particular, and must report to the government the disciplinary measures they have taken so that ministers can decide whether their financial support should be reduced. The edict has been well calculated to madden the liberal English-speaking universities and the radical University of the Western Cape, at all of which the new academic year has been marked by demonstrations of the kind the government wishes to ban, together with the now familiar display of police in riot gear firing tear-gas shells. The next few weeks should show how the five universities concerned will be punished. Most probably, the first penalties will be moderate. But whatever they are, the South African university system will be further damaged. In the long run, the government of South Africa itself will be counted among the losers.

But, for the immediate future, the government will believe it has won a tactical success, in which it will be correct. The new edict has already forced South African universities into two camps. Although the past few years have seen even the Afrikaans universities veer towards opposition of apartheid, the government could well have calculated in advance that they would reluctantly knuckle down to the new edicts. By doing so, they have isolated the more outspoken universities and made them more vulnerable to further threats to their independence. But most of those living in South Africa will see the developing situation in a different light. Surely, a majority of voters will say, the government is asking only that public institutions living on public funds should behave in a seemly fashion, not making a show of biting the hand that feeds them? The government will be able to count on the support of the growing body of opinion in South Africa hostile to the outspoken universities. With its long record of double-speak, the government will soon be pretending that it has erred in not moving against the outspoken universities much sooner.

Damage

Even the immediate damage will be serious. However modest may be the first round of financial penalties, their consequences will be serious. The three largest English-speaking universities (Cape Town, Natal and Witwatersrand) are already in a precarious financial condition because of their commitment to the relatively expensive process of educating non-white students. With their eyes open, they have been skimping on resources for more conventional activities.

The result is that even marginal reductions of the government subsidy will further undermine the well-being of these universities. Is it too much to ask that universities and academics elsewhere, and especially the private foundations in the West (which have already done much to launch non-white higher education), should put their hands in their pockets, however empty, to help sister institutions threatened as are the five front-line universities of South Africa?

Recent correspondence in *Nature* has shown that the academic community elsewhere is divided in its opinion of the role of the outspoken universities of South Africa, but there is sense in which the new situation in South Africa, by threatening the independence of a small group of universities, provides a precedent that threatens all. The sad case of the University of Malta, virtually extinguished a decade ago by a hostile government, should be a warning of the fragility of the university as an institution.

Indeed, there is an uncanny similarity between what is happening in South Africa and the British government's plan to require that British universities should guarantee "free speech" to all visitors to university campuses, and that they should call in the local police if free speech seems threatened.

Opponents

Again, the government's requirement seems only reasonable: why should students be allowed to behave badly at meetings addressed by their political or intellectual opponents? As in South Africa, the government seems merely to be asking that universities should behave in a seemly fashion, which is fair enough. But governments seem congenitally slow to understand the difference between a demand that universities should conduct themselves reasonably and a decision to legislate for reasonable conduct. Robbing a university of its responsibility for the conduct of its own affairs is a means by which its social value is impaired.

That is how South Africa as a whole will in the end suffer if the present threat against outspoken universities is carried out. The turbulence of the outspoken universities is a fair measure of the degree to which the intellectual community in South Africa is opposed to the government's policy on apartheid. But the whole of the university system in South Africa is the government's best hope that some successor will be able to navigate a course to a more just and stable society. The University of Natal has already played an important part in suggesting how novel constitutional forms might help solve the country's glaring problems; even the present frightened government seems to find that development interesting. Academics at the Afrikaans University of Stellenbosch (but not the institution as such) have played an important part in persuading Afrikaans opinion that the encouragement of change is an urgent need.

As the opportunities for peaceful change in South Africa become more remote, politicians there will come to value the work being done by the host of research institutes now working on the definition of what is called the post-apartheid society, and will come to recognize how great a contribution could be made by the present graduates of its sometimes-uppish universities. Even now, there is a clamant need, which the government accepts, to strengthen the infrastructure of non-white education. To punish the universities whose members do not hide their opposition to the government's stand on apartheid will attenuate their capacity to contribute to the solution of commonly recognized problems. That is not the course of rationality, but of folly. Academics elsewhere should say so, and do what they can to help. □



Next step on arms

What should happen now that the treaty on intermediate missiles seems within reach?

HICCOUGHS notwithstanding, it seems that the treaty to abolish land-based missiles of intermediate range (INF) will be signed at some stage later in the year. That seems the burden of the visit two weeks ago to Washington of the Soviet Foreign Minister, Mr Edward Shevardnadze. President Reagan plainly wants the treaty, and will not now have to persuade Mr Caspar Weinberger, who has resigned as US Secretary of Defense, that the treaty almost negotiated at Geneva can be safely signed. It also seems that Mr Mikhail Gorbachev, after a puzzling hesitation (see *Nature* 329, 749; 1987) now also wants the treaty. Politicians will no doubt eventually decide whether Gorbachev's prevarication was a measure of his opposition to the Strategic Defense Initiative or of his doubts about his domestic political support. What matters most is that both sides seem prepared to sign the treaty and confident that it will be ratified, but it is also important to know what will happen next.

Military balance

First, although INF will primarily affect the military balance in Europe, it will not mean that nuclear weapons are abolished from the potential theatre of European conflict; conventional nuclear bombs carried by aircraft and so-called "battlefield" weapons will remain, as will the possibility that strategic weapons based elsewhere might be aimed at European targets. Even so, the INF treaty is bound to stimulate important changes in the disposition of conventional forces on both sides of the fence in Europe. Intermediate missiles now based in Western Europe owe their presence to the decision of the NATO governments at the end of 1979 that nuclear forces should be 'modernized' to meet the threat posed by the then-new force of Soviet SS-20 missiles within striking distance of Western Europe. Now there will be a clamour that conventional forces should be 'modernized', which will dig into the budgets of Western European governments more deeply than giving houseroom to US missiles of intermediate range. That will increase the incentive for seeking some kind of accommodation between East and West on the disposition of conventional forces in Europe. Even the governments of Britain and France, devoted as they are to their independent nuclear forces, will feel the pressure. Moreover, there is now a workable forum (the Helsinki agreements on European security) in which a deal might be struck. That is a project on which both sides should move quickly.

The question of strategic missiles is similarly urgent, if only because INF will be unstable unless quickly complemented by an agreement on strategic arms. What, otherwise, is to prevent the two sides adjusting the patterns of their strategic forces so as, more expensively, to make up for the disappearance of the intermediate missiles? But the prospect that the INF treaty will be signed is a good omen for a strategic agreement, if only because of the provisions it includes for the verification of compliance by on-site inspection of the other side's military and industrial facilities. Although information from satellites is more relevant to the monitoring of a strategic agreement than of INF, it is plainly valuable that INF will be a precedent on inspection that will also be needed for the strategic agreement.

The other side of this coin is, unfortunately, far from smooth. The bilateral negotiations at Geneva have been aimed at a 50 per cent reduction of missiles by each side, but it is not yet agreed whether these reductions should allow for the British and French strategic forces. On the Soviet view, it would be anomalous that these forces should not be counted in the agreement (as they were, implicitly, in the unratified SALT II treaty of 1978). The British and the French, however, will compare themselves with other third nuclear powers such as China (unaffected by the proposed strategic treaty), and will

also argue (if pressed) that their independent forces are a necessary military insurance against future changes of policy by the United States. If President Reagan and Mr Gorbachev seriously hope to sign a strategic agreement next year, they had better see that these awkward questions are raised and thoroughly discussed before too much time has passed.

The second potential difficulty is the Strategic Defense Initiative (SDI) and the related question of the interpretation of the Anti-Ballistic Missile Treaty (ABM). Over the past several years, the United States and the Soviet Union have been daggers drawn on SDI, but there is a sense in which the argument is unreal. It is not necessary to share Reagan's belief that SDI will provide absolute security against strategic attack to know that, even if some future US administration were to take a different view, military researchers would continue to brood about the utility of missile defences. If SDI were formally abandoned in the military programme of the United States, the attention given to it in the past few years has ensured that the genie is out of the bottle that may previously have contained it. Similarly, Soviet protestations that SDI is an abomination do not constitute a proof that the Soviet military have no present ambitions of the same kind. The narrow Soviet interest must be chiefly that it should not be committed to a ruinously expensive timetable set by the United States for the development of more sophisticated missile defences of its own.

The best way of striking a bargain between the declared interests of the two strategic powers is therefore to settle for the strict interpretation of the ABM, allowing research to continue but preventing the testing and deployment of the components of a defensive system. The treaty (also unratified) is ambiguous because its text differs from that of a protocol to the treaty agreed, after signature, by the Soviet Union and the United States. (It may also yet turn out to be contentious — but it should not — that there is nothing in the treaty that would forbid the deployment of an early-warning system based on the use of infrared sensors for the detection of boost-phase missiles, as advertised for SDI.) But the plain truth is that neither side would be believed if it undertook never to engage in military research to explore the feasibility of shooting down the other side's missiles. No treaty can inhibit curiosity, in which sense the question whether SDI as conceived of in the United States would be effective is, strictly speaking, irrelevant. Who can tell that some other system might not be more promising?

Hopes

Is it reasonable to hope that Reagan, committed as he is to SDI, will be willing to settle for a reaffirmation of the ABM in its strict interpretation? Although the crisis on Wall Street has forced him to go back on his unwillingness to increase taxes, there is no reason to believe that he would change his stance on SDI. But not all is lost. Although there were many who were hoping, a year ago, that it would be possible to make a start on the deployment of SDI before the end of the present US administration at the beginning of 1989, that now seems improbable. For one thing, the US Congress has been too niggardly about funds for SDI, but the sheer shortage of launching facilities for Earth satellites precludes all but a token — and inflammatory — installation of anti-missile devices in the next years.

So Reagan could agree to, say, a five-year extension of ABM in its strict form without having to swallow his pride in SDI, leaving his successors to decide what should be done thereafter. (It would be interesting to know whether Weinberger's resignation was prompted by such a prospect.) Gorbachev would almost certainly complain that a deal on strategic weapons could not be struck without a more permanent interdiction of SDI, but he should nevertheless be willing to settle for something along these lines. There would still be many missiles surplus to reasonable requirements, while a breathing space of even a few years would allow the momentum of arms control further to reduce the chance that they would ever be used. □

Patent Office decision puts Genentech out in front

- Broad claims of biotech company granted
- Courts will decide ultimate value

Washington

A BROAD patent covering molecular biological tools used to create proteins from genetically engineered organisms was granted to Genentech last week by the US Patent and Trademark Office. The patent encompasses techniques used by the entire biotechnology industry, and may entitle Genentech to collect royalties from every company selling products made by genetic engineering.

Genentech's patent is being compared to the Cohen-Boyer patent that launched the biotechnology industry. The Cohen-Boyer patent covered techniques of gene splicing and expression of certain genes that marked transformants. Genentech's patent is subservient to the Cohen-Boyer patent, so anyone seeking to license from Genentech will first require a licence from the University of California and Stanford University, the patent's co-holders. Since the Cohen-Boyer patent was awarded in 1980, the universities have received approximately \$6 million in royalties.

The patent includes 15 claims; some exceedingly specific, and others sweepingly broad. The first and broadest claim establishes Genentech's proprietary right over any "recombinant DNA cloning vehicle suitable for the transformation of a microbial host" that consists of a control region regulating the expression of a structural gene, where the gene coding for the polypeptide is in the correct reading frame for expression, and the polypeptide is in recoverable form. The other 14 claims cover specific plasmids used as cloning vehicles, including those containing the *Escherichia coli lac* and tryptophan operon-promoter systems. Genentech also lays claim to plasmids that produce mammalian hormones or polypeptides in general, and specifically ones that produce proinsulin, growth hormone, and the A and B chains of human insulin.

Genentech filed for the patent in 1977, when the enormous commercial potential of biotechnology was just being recognized. The patent office delayed its decision on Genentech's patent until the 1980 Supreme Court decision in *Diamond vs Chakrabarty* cleared the way for the patenting of living things produced by human intervention. The next seven years, according to the company, were spent establishing the priority of some of the claims over work done at the University of California.

In announcing the patent, Genentech's

Thomas Kiley said the company would pursue an open nonexclusive licensing policy. He indicated that it was too early to tell what the royalty rate would be, and that the company would wait until the companies using the technology had sufficient time to approach Genentech for a licence before prosecuting infringers. Genentech will only be eligible for royalties from products after 3 November 1987, the issue date of the patent.

But some say that the patent is not that all-encompassing because it makes no mention of the increasingly popular mammalian cell expression systems, and may not even cover some special microbial systems. Patent attorneys point out that the proof of a patent's value does not come until it is challenged in court. Just because the patent does not expressly cover mammalian expression systems, the courts may nonetheless rule that it is applicable to such systems.

Attorney Jorge Goldstein, of Saidman, Sterne, Kessler and Goldstein, says Genentech could insulate its broader claims from scrutiny by using one of the more specific claims of the patent. In this scheme, a firm sued for infringing one of the more specific claims could not argue against more than the one claim in its suit.

Most of the smaller biotechnology companies are expected to come forward and seek licences from Genentech. It is in the biotech industry's best interest to be pro-patent while the industry is just getting started, and small companies will not want to risk being put out of business trying to oppose a patent that they can license easily. Potential opposition could come from larger pharmaceutical firms with in-house biotech programmes, such as Eli Lilly. According to a Lilly spokesperson, Lilly "will be unaffected by the [Genentech] patent", even though Lilly markets a competitor to Genentech's human growth hormone that is also produced using recombinant DNA techniques. Lilly would have the resources to contest the patent, but legally would have to wait until it received an infringement suit, or at least a letter threatening a suit, before it could mobilize. Most of Genentech's products, both on the market and under development, are also being worked on by other companies. Genentech could use the patent to lock out competitors, and potentially even competitors with protein-engineered or "second generation" products.

Carol Ezzell

New radiotelescopes

THE University of California at Berkeley is to add 3 six-metre radiotelescopes to its Hat Creek Radio Observatory near Mt Lassen in northern California.

The \$3.6-million project, a joint venture of the Universities of California, of Maryland and of Illinois, will increase the data collection capabilities of the observatory fivefold. Data from Hat Creek will be fed to powerful new computer facilities at Illinois and Maryland.

M.B.

Einstein for sale

A HITHERTO unknown manuscript on relativity, written by Einstein in 1912, will be auctioned at Sotheby's in New York on 2 December. At 72 pages it is one of the



longest and clearest expositions of his theory and casts new light on the development of his thinking. Bids of up to \$700,000 are expected for the manuscript.

Written in German, it was intended as a chapter for the *Handbuch der Radiologie* that Professor Erich Marx of Leipzig published in five volumes between 1913 and 1925. Unpublished correspondence between Marx and Einstein suggests that its publication was halted by the First World War.

S.J.H.

Biotech expansion

THE European Commission is calling for a £14-million expansion of its biotechnology programme to update information systems, improve training and make room for new European Community members Spain and Portugal. The present £38 million programme expires in 1989. A five-year follow-on programme is scheduled to start in 1990.

B.C.

UNESCO's new head

THE appointment of Spanish molecular biologist Frederico Mayor as the director general of UNESCO was formally ratified on 7 November. At a full assembly of UNESCO member states in Paris, Mayor received 142 votes out of 149, to replace the controversial Amadou Mahtar M'Bow of Senegal (see *Nature* 329, 659; 1987).

S.L.H.

Castration pleases bulls?

New Delhi

A VACCINE that sterilizes bulls, developed by biotechnologists at the National Institute of Immunology (NII) at New Delhi, is to be used as a key tool in India's campaign to boost milk production.

Tests carried out at government farms have shown that bulls become 100 per cent sterile within four weeks of a single injection into the testis. The sterility produced is said to be permanent.

India's programme for cattle improvement rests primarily on artificial insemination of cows and buffalos using semen of high-quality exotic bulls. Last year alone, nine million cows were inseminated artificially at India's network of 15 semen-freezing stations, 52 frozen-semen banks, and 5,500 artificial insemination centres. But the AI programme has so far failed to make impact, one major reason being that the cows are impregnated by local bulls before the arrival of semen from semen banks. The NII vaccine will now be used to sterilize these interfering bulls. The cost of sterilizing an animal is said to be less than one rupee.

Trials have begun in Bhoosali, a village in Haryana state bordering Delhi. Local bulls in the village have been sterilized with the injectible vaccine, and the National Dairy Research Institute (NDRI) is planning to extend the work to 30 more villages in the state. If the experiments are useful, sterilization will be taken up on a national scale. The idea, according to an NDRI spokesman, is gradually to improve the stock of Indian cattle. Out of the estimated 180 million cows and 64 million buffalos in the country, two-thirds are of non-descript quality.

Traditionally, Indian villagers used to sterilize poor quality bulls by mechanically crushing the spermatic cord. Apart from being cruel and painful, this method also led to total loss of libido thereby depriving the bulls of their ability to detect females on heat. The vaccine has no such effect. For AI to be successful, it must be done within 48 hours of ovulation and bulls, rather than veterinarians, are the best detectors of this crucial period.

The NII has not disclosed the composition of the vaccine except that it is based on BCG (Bacillus Calmette-Guerin) used for tuberculosis prevention. It relies on eliciting an immune response at the site of sperm production. Trials in rats, guinea pigs, rabbits, dogs and monkeys had shown that BCG injected into the testis could block spermatogenesis without the loss of androgens. According to NII, the technique works in all mammals without producing any adverse side-effects.

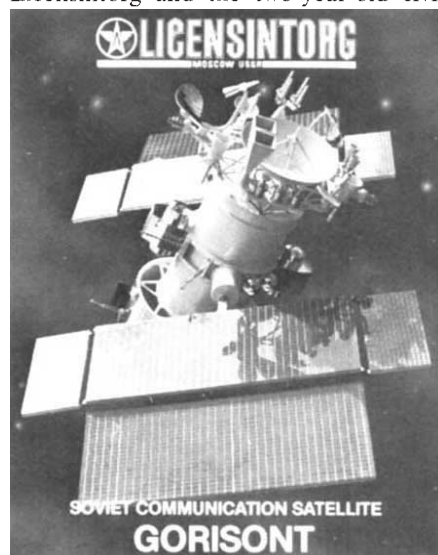
K.S. Jayaraman

Soviet space agency seeks first foothold in tough Western market

London

THE Soviet Union seems to have made its furthest progress yet in its attempt to break into the lucrative market of commercial space launches with the signing of preliminary agreements with three Western companies. But the initiative is unlikely to proceed further unless the United States relaxes its virtual embargo on the transport of US technology to the Eastern bloc — almost all Western-made satellites contain components from the United States.

The latest efforts by the Soviets to force a change in the rules were disclosed in London last week when representatives from the Soviet foreign trade organization Licensintorg and the two-year-old civil



Hard sell — a glossy brochure Soviet-style

space agency Glavkosmos briefed British industrialists on the full range of services available. London-based Jardine Insurance Brokers has been signed up "to market Soviet launch vehicles to commercial communications companies in Europe and newly industrialized countries".

There is little dispute that the Soviets are offering a commercially attractive deal — a two-tonne payload can be placed in geostationary orbit for between £18 million and £20 million, launched by the reliable Proton vehicle (with an impressive record of only nine failures from 104 launches between 1970 and 1987), representing an undercutting of Western rates by between one-half and one-third. Similarly favourable insurance arrangements are available through the Insurance Company of the USSR (Ingosstrakh) Ltd.

The Soviets are keen to dispel any fears about the security of payloads. The launch package enables the customer to accompany the payload at all stages, from chartering an Aeroflot flight from the

country of origin direct to the Baikonur launch site (waiving normal customs regulations), to seeing the equipment on board the launch vehicle. Much is made of the facilities available to foreign technicians escorting the payloads, including "comfortable rooms, restaurant, wine bar, swimming pool and sauna".

Maurice Brackenreed Johnston, deputy chairman of Jardine, seems confident that the three companies who have signed preliminary agreements will have their satellites in orbit within three years. "We went into this with the full knowledge of all the difficulties we were going to encounter and the length of time it was going to take us to develop this relationship", Johnston said last week.

The chief obstacle to the Soviet Union's entry into the global launching business comes from the United States International Traffic in Arms regulations, section 126.1, which specifically forbids the transport to the Soviet Union of any technology of US origin without prior agreement from the State Department, which has made it abundantly clear in the recent past that such permission will be refused.

Within the past 18 months the Soviet government has made official representations to the White House urging a relaxation of the rules and Glavkosmos has lobbied satellite manufacturers anxious to clear their mounting backlog ("dangling fat carrots before their eyes and waiting to see what the outcome would be", according to one sceptical State Department source). So far the outcome has been negative and officials see little indication of any change in the near future.

Furthermore, it is suspected that the Soviet commercial launch programme would be heavily subsidized. Last week the Soviet delegation in London was obviously aware that this thought was likely to be occurring to people at the meeting; without prompting, its chief spokesman, Evgeni Louppov, a senior engineer with Licensintorg, said how an agreement had not been signed with one Western company which was not prepared to pay the rate required for Glavkosmos to realize a sufficient profit. For the Soviets — and Jardine — the US regulations have little to do with the transfer of technology and a lot to do with protecting the interests of Western launch vehicle manufacturers.

As launch orders continue to accumulate in the West, political pressure on the US government to let the Soviet Union into the launch market is certain to increase. Just how long it will be before the wine bar at Baikonur serves its first party of Western technicians remains to be seen.

Simon Hadlington

New effort starts up to evaluate AIDS education in United States

San Francisco

THE largest effort to date to study the effectiveness of AIDS education in preventing infection by the virus is being funded by the National Institute of Mental Health (NIMH) and the National Institute of Drug Abuse (NIDA) at three centres around the country. Those who have been calling out for more research into AIDS prevention will say that the initiative is long overdue.

Although public health officials agree that education programmes are virtually the only tool for combating the spread of HIV (human immunodeficiency virus) infection, evaluating these programmes is just getting under way. Stephen B. Hulley, director of the Center for AIDS Prevention Studies at the University of California at San Francisco (UCSF), says the new centres represent the "first sizeable efforts in the whole field" of research on AIDS education and prevention.

The UCSF centre is a collaboration between the university, the San Francisco Department of Public Health, and MIRA (Multicultural Inquiry and Research on AIDS), a group of researchers who study the impact of AIDS on minorities. Funding comes from a recently awarded \$7.1-million, four-year grant, in addition to the \$3-million, five-year grant that launched the centre last year.

Research at the UCSF centre will include improved epidemiological studies to learn the prevalence of high-risk behaviour, and research to test the effectiveness of educational programmes in reducing that behaviour. The centre's 50 investigators will focus on intravenous drug users and their families, gay and bisexual men, and ethnic minorities. The centre is also conducting a study of high-risk behaviours in Rwanda, Africa and an international exchange programme to help researchers from Africa and Latin America study AIDS prevention in their own countries.

Hulley says the MIRA researchers are integral to the centre's programme, a major aim of which is to investigate the factors behind the disproportionate representation of minorities among people with AIDS.

Black and Hispanic researchers will have better cultural understanding of their minority study subjects, says Hulley. MIRA and health department researchers will provide contact with existing AIDS education programmes, facilitating the rapid translation of research results into more effective programmes and policy.

A second centre is Columbia University's HIV Center for Clinical and Behavioral Studies in New York, which

received a \$19-million award last month. The New York Psychiatric Institute and Columbia Presbyterian Hospital are partners with Columbia University in the centre, and other regional hospitals will participate in the centre's activities.

Three of the five studies proposed by the centre will analyse the effectiveness of AIDS prevention education, with the emphasis on how best to reach adolescents, said Claudia Bial, spokeswoman for the centre.

Although members of the New York health department will not participate directly in the research, Bial says a close cooperation with the city will allow the researchers access to the adolescent high school students, runaways and sex-offenders who will be the subjects of the studies.

A third AIDS centre at the University of Miami School of Medicine received a \$5.5-million, five-year grant from NIMH and NIDA in November 1986. Called the Biopsychosocial Center for the Study of AIDS, the Miami centre is focusing its attention on the development of AIDS dementia and on how social and psychological factors influence the course of HIV infection.

Marcia Barinaga

Fewer projects funded in West Germany

Munich

THE interdisciplinary West German research programmes called Schwerpunktprogramme have been reduced this year, reflecting the difficult research funding situation. The Deutsche Forschungsgemeinschaft (DFG) decided in Bonn on 15 October that it would fund only eleven out of twenty-four applicants, down from the usual acceptance rate of two-thirds. The programmes run for five years and support research projects carried on in several laboratories at once, with yearly meetings among all involved researchers.

Schwerpunktprogramme account for almost 20 per cent of the research support given by the DFG. This year's new programmes focus on biosciences. Included are projects titled "Molecular and immunological mechanisms of host-parasite interactions" and "Chemical ecology". Projects were also approved in humanities and social sciences, natural sciences and engineering. The 110 Schwerpunktprogramme funded by the DFG will receive DM 188 million in 1988. The overall budget has remained roughly constant since 1985. Steven Dickman

First images from European telescope



THE picture shows the Roque de los Muchachos observatory on the Canary Island of La Palma, home of the William Herschel telescope, which recently started yielding its first images. The telescope, which took four years to build at a cost of £12 million, is run by Britain's Royal

Greenwich Observatory in collaboration with the Netherlands Organization for the Advancement of Pure Research. It has a 4.2-metre mirror, making it the world's third largest single-mirror optical telescope after the US Palomar 5-metre and the Soviet 6-metre telescope. □

Patience pays laser inventor in lifetime battle over patents?

Washington

A 28-year legal battle ended last week with Dr Gordon Gould's claim to have invented — or at least improved — the laser finally vindicated. On Wednesday he was presented with a patent for the gas discharge laser that he had applied for in 1959 and has been fighting for ever since. Ironically, Gould is far better off for having been forced to wait: had he received his patent quickly it would have expired long before the laser industry took off — now the patent runs until 2004 and may be worth hundreds of millions of dollars.

Last week's success follows hard on the heels of victory in a legal fight over an-

Pending patent law good news for some

London

THE British government has finally published details of long-awaited legislation to get to grips with outdated and confused copyright and patent laws. The Copyright, Designs and Patents Bill is scheduled to be put before the House of Lords this week, and is expected to become law by next summer.

One prominent feature of the bill is that the government has resisted pressure from the music industry by backing down from imposing a 10 per cent levy on blank audio tapes to compensate copyright holders for lost revenue from people recording tapes at home. Plans to impose a levy had been announced in the policy document that preceded the bill (see *Nature* 327, 549; 1987). Home taping and video taping will remain illegal, despite the government's admission that the law is virtually unenforceable.

The bill brings good news for the pharmaceutical industry by removing the 'licence to right' that allows companies to copy successful drugs during the last four years of their 20-year patent protection. The government says the modification of the law is in recognition of the length of time it takes from patenting a new drug to its appearance in the market place, which is often as long as 12 years, and which gives companies little time to make profits.

The new law will make it easier and less expensive for disputes over patents and designs to be settled by allowing actions to be brought before a lower court than is presently the case, and the problem of getting copyright clearance to record broadcasts for use in school is removed.

Simon Hadlington

other key patent — for the optically pumped laser — that is also held by Gould. Although that patent was issued in 1977 it remained under appeal at the US Patent Office for almost ten years. Only two weeks ago was it finally validated when a patent infringement suit was won against Control Laser Corporation in a Florida court. Gould's two patents together could cover as much as 80 per cent of the lasers being produced in the United States — a market worth \$500 million a year now and doubling in size every four years.

Sharing Dr Gould's obvious happiness were the officers of Patlex Corporation, a company set up specially to help Gould and which has spent millions of dollars backing his claims in the courts. That has been a considerable gamble, for the company's biggest asset is its 64 per cent income interest in Gould's patents. Elsewhere, the company part owns three small Israeli high-technology armament manufacturers and a US laser company.

Despite legal recognition of his claim to have invented the laser, Gould has yet to see his name written on the pages of scientific history. He says he set down the ideas for the laser in two straight days of thinking and writing when he was a graduate student at Columbia University's Radiation Laboratory in 1957. Columbia, he says, was a "terrific place . . . the foremost laboratory in the world in microwave spectroscopy, the field in which the laser had to be born". Gould had already received training in optical spectroscopy at Yale University and that, in combination with the microwave techniques available at Columbia, led him towards the laser. But it was not until two years after he set down his ideas of the laser, in 1959, that he applied for a patent. In the interim, a patent application for the laser, then known as the "optical maser" had already been filed by another Columbia University researcher, professor Charles Townes.

Townes has always been regarded as the father of the laser. His paper, published in *Physical Review* in 1958 with A.L. Schawlow, is seen as a milestone in the history of the laser. And in 1964 Townes won the Nobel prize, together with two Soviet scientists, for the fundamental work in quantum studies that made the maser and laser possible.

Patlex officials, with great pride in their protégé, believe that Gould should also have been a candidate for the Nobel Prize. But Gould himself is more sanguine. He says only that it was all his fault that he was not recognized for he was "too ignorant" to know how to apply for a patent when he

first had the idea for the laser. He sees his own and Townes's work as complimentary. Townes's and Schawlow's classic paper showed that the Fabry-Perot interferometer provided the appropriate cavity for laser oscillation. Gould built on Townes's studies but concentrated on the appropriate excitation mechanisms.

There was no close link, Gould says, between Townes and himself because "I was a student and he was a professor". Townes remembers Gould well as a student, but regards his contribution as more of an improvement than the addition of something new. Gould, in any case, appears little concerned with the events of



Gordon Gould — in the money?

the past. He is much more worried about the extraordinary difficulties of patent licensing in the United States and the effect this may have on individual inventors and the ability of the United States to compete with other countries. Nor are an inventor's worries over when a patent has been granted. As Gould puts it "all a patent gives you is the license to sue infringers". Both legal action and examination of a patent can be ruinously expensive and a patent can be overturned in court after it is issued.

Gould believes his patent will survive. In the decades it has taken to obtain his patents every possible legal objection has already been raised, he believes. In the United States, at least, he thinks that companies are now more likely to enter licensing agreements than to go to the expense of challenging the patent in court. But an unpleasant surprise may yet come along. The patent will not apply abroad for, with the principal exception of Canada, patents run from the date of application rather than the date of issue and, in Gould's case, would long since have expired. But as the United States is by far the largest producer of lasers, Gould may be able to sit back and see the millions flow in.

Alun Anderson

Extra money for British science insufficient to reverse decline

London

NEXT year's increase of £47 million in the science budget over previous plans will provide little respite for British academic science. The increase, announced last week by Education Secretary Kenneth Baker following the Chancellor's Autumn statement, falls well short of the £103 million deemed necessary by the Advisory Board for the Research Councils (ABRC). Of the extra money, which will bring the total science budget to £696 million, £28 million will pay for agreed salary rises, £6 million will go to a directed programme of research into AIDS and £8 million to the Natural Environment Research Council (NERC) towards a replacement ship for the British Antarctic Survey. The planning figure for each of the following two years is now £729 million, an 11 per cent increase over this year's budget.

The ABRC will decide in the next few weeks how the money is to be divided among the research councils. Initial reaction to Baker's announcement vary. Science bodies deplore the increase as derisory, with the lobby group Save British Science claiming that after accounting for inflation and correcting the sterling exchange rates on international subscriptions, a net decrease of £19 million will result.

For the administrators in the research councils, the money will probably enable the books to be balanced, although highly rated research projects will continue to be shelved and no money will be available to reverse earlier decisions to cut back on areas of research.

In its annual advice to the Secretary of State, the ABRC recommended an increase in the science budget of £32 million for 'necessary underpinning', £27 million for the replacement of obsolete equipment and £44 million for the reshaping of the science base (*Nature* 328, 369; 1987). However, the recommendations built on previously submitted advice on reorganizing academic science, contained in the document *A Strategy for the Science Base* (*Nature* 328, 280; 1987). There seems to be some confusion about whether the strategy advice was considered when finalizing the budget. On the one hand Baker stated that the extra money would give scope for a "worthwhile start in 1988–89 on the establishment of interdisciplinary research centres, as recommended by the ABRC and to which I attach great importance". However, he then said that the government would be giving further consideration "in the next few months" for the strategic reshaping of the science base, taking account of ABRC's strategy advice and subsequent consultations.

One of the first things the research councils will need to know is whether the ABRC will be able to confirm the £12 million of provisional allocations it made some months ago using its 'flexibility margin'. For the Science and Engineering Research Council (SERC), which receives more than 50 per cent of the total budget, there should be money available for it to meet its major commitments — salaries, international subscriptions and the operation of its recently installed Cray computer.

The number of research programmes SERC will be able to support under the government's LINK scheme (which encourages collaboration between industry and universities), and whether Britain will be able to afford to subscribe to the European Synchrotron Radiation Facility, will be decided in December. Plans to establish interdisciplinary research centres will probably continue. A centre for superconductor research is being set up, and two to four new centres will be established next year depending on the precise allocation of funds between the research councils.

Eyebrows have been raised over the allocation of £8 million to NERC towards a replacement for the British Antarctic Survey's 30-year-old vessel *John Briscoe*. In its strategy document, ABRC suggested that NERC's oceanography programme, with its associated fleet of vessels, is an area of big science "which we think may call for review". According to Jeremy Bray, the opposition Labour Party's spokesman on science, "presumably the new ship is needed to buttress the government's 'Fortress Falklands' policy". Privately, NERC officials concede that political interest in the Antarctic could have played a role in the decision to replace the vessel.

The Autumn statement provided for an extra £56 million for the universities for the 1987 academic pay settlement. Baker says the money will be released "as soon as the government is satisfied that the universities have sufficiently improved their arrangement for staff appraisal, probation and promotion". A further £61 million is to be made available in 1988–89 and later, £53 million of which will be earmarked for 'targeted programmes' of restructuring, including additional money for early retirements, 'new blood' appointments; 'rationalization' of a number of disciplines and an initiative in manufacturing systems engineering. Polytechnics and colleges will receive an extra £64.5 million, with an additional £9 million to help pay for their pending transfer from local authority control.

Simon Hadlington

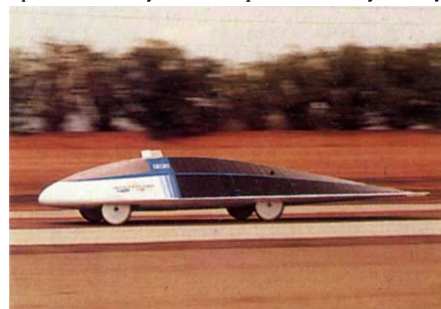
Flying cockroach walks to victory

Sydney

THE winner of the world's first trans-continental solar-powered road race from Darwin to Adelaide in Australia crossed the finish line last week, days ahead of its closest rival. An odd assembly of 24 solar vehicles which had gathered from around the world started on a 3,004-km trek from the north coast through Australia's desert heart to Adelaide in the south. When Sunraycer — the General Motors entry that completed the course in five and a half days — crossed the finish line, some of the other entrants had yet to reach Alice Springs.

Sunraycer's dominance of the race was obvious within 10 km of the starting line. Shaped like a flattened teardrop, and called the "Flying Cockroach", it can sustain speeds around 110 km h⁻¹. Its average speed over the course was 66.7 km h⁻¹.

High technology was the key to Sunraycer's success. The expertise in low-speed aerodynamics provided by body



designer Paul McCready gave Sunraycer an extremely low drag coefficient as well as relative impunity to cross winds that caused headaches for other competitors. It used gallium-arsenide solar cells rather than the silicon cells used by the rest of the field. Sunraycer's 92 per cent efficient electric engine contained magnets made of an alloy of the rare earth neodymium. The alloy has an amorphous crystal structure, like that of glass, created by cooling the molten alloy so quickly that crystals have no time to form.

The energy may be free but the cars certainly are not: the Sunraycer cost between \$3 million and \$15 million.

The rules of the race limit each car to 8 m² of solar cells. The cars race from 9 a.m. to 5 p.m. and park for the night at whatever point they have reached at that time. Cars which have reached Adelaide within 5 days of the winner are officially classified as having finished the race.

The third-place Swiss entry, Spirit of Biel, lost second position after a collision with a local motorist in Alice Springs. The police are to issue an infringement notice to the Swiss for failing to give way.

Charles Morgan

United States and Japan sign agreement on nuclear reprocessing

Tokyo

NUCLEAR power entered a new era in Japan last week with the signing of an agreement with the United States. The new bilateral pact gives Japan greater freedom to reprocess spent nuclear fuel, opens the way to commercial reprocessing in the mid-1990s, and allows the export of Japanese nuclear power technology to the United States for non-military purposes. But the agreement also calls for the shipment of vast quantities of plutonium to Japan by air after reprocessing in Europe. And there is strong opposition to the plan in Alaska, which lies under the most likely flight path.

Despite the accident at Chernobyl, the Japanese government is eager to expand

its nuclear power facilities. The Atomic Energy Commission forecast earlier this year that by the year 2000 nuclear power will provide 40 per cent of the nation's energy, compared with the present level of 27 per cent, and 7 million million yen (\$50,000 million) will be invested in the nuclear power industry over the next 13 years.

Japan already has more than 30 nuclear power stations in operation, which produce around 700 tons of spent fuel a year, but only about 10 per cent of this can be handled by Japan's small reprocessing plant in Tokaimura, Ibaraki Prefecture; the bulk of the nation's spent fuel is reprocessed in Britain and France.

Under terms of an agreement signed in

1968, Japan must obtain prior approval from the United States before shipment or reprocessing of any fuel of US origin (a condition which applies to most of Japan's nuclear fuel).

The new pact, signed in Tokyo last week by Japan's foreign minister and US ambassador Mike Mansfield, grants Japan blanket approval for the reprocessing and transport of recycled fuel for the next 30 years, a move that should allow the smooth operation of a commercial reprocessing plant due to open in Aomori Prefecture in 1995 (see *Nature* 322, 399; 1986).

During the late 1970s the Carter administration in the United States, which advocated a strong policy of nuclear non-proliferation, caused considerable delays to the opening of the Tokaimura plant, and it was largely to break free of such restrictions that Japan began negotiating the present pact with the Reagan administration in 1982.

But even after the plant in Aomori opens, Japan will continue to send vast amounts of spent fuel to France and Britain for reprocessing. The new accord requires that the plutonium produced in Europe should be shipped back to Japan by air. With an expected 40–50 tons of plutonium to be transported in the 1990s, this would require several flights a year. The expected route is over the North Pole and Canada, with a refuelling stop in Anchorage.

Alaskan governor Steven Cowpar has considerable reservations about the plan and has asked Congress to seek necessary environmental safeguards before approving the new pact. In particular, he demands assurances that the plutonium will be shipped in crash-proof containers. But, as pointed out by the US Nuclear Control Institute in a report last year, suitable casks have yet to be developed. And there are fears that in the event of a plane crash the highly toxic plutonium could be sprayed across the countryside.

There are also concerns that Japan may not take adequate measures to prevent hijacks. A few years ago the United States initially vetoed a plan by Japan to transport a large plutonium load from France by ship because of inadequate preparations for escort and surveillance. But critics say that with the signing of the pact such veto rights are lost.

So what does the United States gain out of this agreement? Japan has made considerable advances in the development of technology for the nuclear power industry, for example, advanced technology and robots are being used to dismantle Japan's first experimental reactor (see *Nature* 325, 100; 1987). And under the accord, such technology can be transferred or exported to the United States, provided it is used for non-military purposes.

David Swinbanks

Axe falls on Irish agricultural researchers

DESPITE recent declarations of the economic importance of increased research and development, the Irish government is to reduce its support of agricultural and food research by almost half next year. An Foras Taluntais (AFT), the agriculture and food research institute, is to be merged with ACOT, the agricultural advisory and training service. The state grant for the new body will be IR£20 million (about £22 million), a 43 per cent reduction on the joint total of IR£35 million this year. Voluntary early retirement and redeployment of staff within the public sector is expected to achieve a staff reduction from 2,200 to 1,200.

Although rumours of a merger have been circulating for some months, the severity of the financial cuts has taken everyone by surprise, particularly as the government had been emphasizing increased research and technological development as a key factor in its "programme for national recovery", announced last month. The cuts are also surprising given the importance of agriculture to the Irish economy.

Almost 12 per cent of Irish gross national product (GNP) is attributable to the agricultural sector, compared with 2–4 per cent of GNP for most other members of the European Economic Community (EEC). However, proportionally less state money is invested in agricultural research in Ireland than is the case in most other developed countries. In 1985 the Irish government spent IR£178 per agricultural worker; on a comparable scale, Belgium and The Netherlands spent twice as much, whereas Britain spent three times as much.

AFT's state grant this year is IR£17 million, 67 per cent of its total income, the balance coming from farm operations and

research contracts (with more than IR£1 million from EEC projects). Even before the latest cuts take effect, the state grant has declined in real terms by 25 per cent since 1981. Over the past five years, nine field stations have been closed and staff members reduced by 200 to the present level of 1,134.

The Department of Agriculture, which administers the AFT and ACOT grants, says the cuts are necessary "to reduce public expenditure". Within the department, however, AFT and ACOT seem to have been singled out for particularly drastic reductions — the department's allocation for its own administrative budget next year is only 2 per cent down on this year.

The severity of the cuts will almost certainly mean the closure of whole areas of research. So far the government has not indicated its priorities, and it now seems that the newly merged body will decide at a local level how best to implement the reductions.

This apparent lack of any underlying strategy by the government has caused considerable concern to AFT's 240 research staff who now fear that the government is forgetting the primary role played by research and the logistics involved in running a comprehensive research programme.

The cuts to AFT and ACOT are only the latest in a series of government measures to reduce public expenditure. The Institute for Industrial Research and Standards and the National Board for Science and Technology, which together employs 650 staff, are to be merged with a loss of 100 jobs, and An Foras Forbartha, which employs 200 people and is responsible for physical and environmental planning, is to be abolished.

Mary Mulvihill

University of California will stay in charge at weapons labs

Berkeley

FOR the University of California (UC), operating the Department of Energy (DoE) weapons laboratories at Los Alamos, New Mexico and Livermore, California has been a public relations nightmare. And with the expiration of the law that protected UC from liability claims in the event of an accident at one of the laboratories, operating them threatened to become a financial nightmare as well. But UC Regents believe the service is worth the trouble.

For years, critics have charged that it is inappropriate for a university to be running top-secret facilities primarily devoted to designing and testing nuclear weapons. But William Frazer, UC vice-president, says if it is national policy to have such facilities, they should be run in as open a manner as possible, something a university is best equipped to do. Despite criticisms, the University decided two years ago that, when the existing 5-year contract with DoE expired, it would seek a renewal.

But that was also two years before the expiry date of the Price Anderson Act. That law was enacted to provide liability protection for the nuclear power industry in the event of an accident at a power plant. But it also contained provisions that indemnified DoE contractors. DoE not only owns its own research reactors, it is also the agency responsible for building nuclear weapons. Contractors say that liability protection is essential to make operating such facilities financially viable.

Last year it seemed that a compromise had been reached permitting the renewal of Price Anderson and raising the limits of liability for nuclear plant operators to levels in line with the cost of an accident (see *Nature* 322, 676; 1986). But at the eleventh hour, the compromise fell apart, and renewal legislation had to be reintroduced.

For the university, this posed a dilemma. Its contract with DoE expired on 30 Sep-

tember, and without liability protection, UC felt the financial risk was too great to renew the Los Alamos and Lawrence Livermore contracts. So DoE came up with the idea of indemnifying its contractors under Public Law 85-804, part of the War Powers Act. Unlike Price Anderson, the War Powers Act does have a condition in which UC could be liable for damages. But it would take an act of wilful misconduct or bad faith by the laboratory direc-



Frazer — being open about weapons labs.

tor or one of his superiors to make UC liable, and then liability extends only to government property. With the guarantees provided by 85-804, UC and DoE agreed a new five-year contract on 18 September.

But UC officials say the new coverage is a bad deal for the public. UC counsel Glenn Woods explained that the War Powers Act coverage has less protection for claimants seeking damages under Price Anderson.

Other DoE contractors are not sanguine about liability coverage under the War Powers Act. Several say they will not sign future contracts without Price Anderson-type protection. Last month DuPont announced that it would not renew its contract to operate DoE's Savannah River plant where plutonium and tritium are

into the continent, ocean currents pick them up and they begin to circulate eastwards in the circumpolar flow. Eventually they break up and melt, but large icebergs can survive for many months.

Icebergs in the Southern Hemisphere are not the same threat to shipping as in the north; the icebergs may drift as far in latitude, but southern shipping lanes are much further from the polar regions. The only danger might be to scientific survey ships, but the US Navy and the National Oceanic and Atmospheric Administration will be watching the iceberg's progress carefully.

David Lindley

Efficiency bid by research commission

Brussels

AS part of a determined effort to increase efficiency in its Joint Research Centre (JRC), which runs laboratories at four separate sites, the European Commission is proposing to split it into nine autonomous institutions. It also wants to abolish the scientific council that oversees the work of the JRC so scientists will have a greater say in the research programmes, and to answer criticisms from members of the European Community (*Nature* 329, 576; 1987).

A new institute for advanced materials will be established at Petten, in The Netherlands. The Central Bureau for Nuclear Measurements at Geel, Belgium, and the Institute for Transuranium Elements at Karlsruhe West Germany, will remain. The Ispra establishment in Italy will be split into five institutes for engineering, environmental sciences, remote sensing, safety technology, electronics, information technology, and a new European Telecommunications Standards Institute. Research ministers meet this month to discuss the proposals.

Bernard Conlon

produced. DuPont said its decision was based in part on the belief that a growing number of elected officials "favour elimination of protection of liability" crucial for its financial security.

Price Anderson's future in Congress is still uncertain. The House of Representatives has passed one version, but it has yet to reach the Senate.

Joseph Palca

● The recently signed contract between UC and DoE for running the Los Alamos, Livermore and Lawrence Berkeley laboratories has some sweeteners for the university. The fee for operating the facilities will climb from approximately \$7 million to \$12 million per year, still a tiny fraction of the nearly \$2,000 million annual budget for the three facilities. The new money will allow the expansion of university-laboratory collaborations.

UC has argued that DoE has been too restrictive in its approach to allowing patents to vest with contractors at its national laboratories. William DeGarmo, counsel for Lawrence Livermore National Laboratory, says DoE has made it extremely difficult for the laboratory to enter into any commercial arrangements with industry. Instead, he says, DoE has put non-classified inventions and software into the public domain, discouraging industry from investing in it. DeGarmo says commercial development of inventions developed with government money has been a time-honoured and successful way of bringing new technology to the market place. "If DoE had developed the airplane, we wouldn't have commercial aviation", he says.

J.P.

Huge iceberg adrift in Antarctic seas

Washington

A CHANGE in the maps of Antarctica will be necessary following the departure of an enormous iceberg, 2,500 square miles in area, from the Ross Ice Shelf. This single iceberg contains two or three times as much ice than breaks off from the Antarctic continent during a typical year. The Bay of Whales, the starting point of many early expeditions, has largely disappeared.

The creation and fate of icebergs is not well understood. As glaciers such as the Ross Ice Shelf run down to the sea, they fragment, and the pieces drift westwards along the coastline. Unless they crash back

Problems of university of Chile

SIR—We have recently received distressing news from colleagues at the University of Chile, Santiago. Over the past ten years, both the university budget and student numbers have been reduced by half and no positions have been opened for the recruitment of young scientists and teachers.

Last September, General Pinochet designated Mr José Federici as rector of the university without any prior consultation with the university council or the professors. This procedure was resented, and 11 out of 13 members of the university council (among them the elected faculty deans) considered that Mr Federici had no credentials to rule the university.

In response to this reaction of the council, Mr Federici dismissed the deans of four faculties: Professors Mario Mosquera (law), Atiliano Alamana (engineering), Fernando Valenzuela

(philosophy) and Hernan Montecinos (architecture). Moreover, Mr Federici closed the university, impeding access by the professors, students, researchers and technicians to the university premises, thus interrupting the development of research programmes in progress. More recently, Mr Federici dismissed 35 other professors from various faculties, some of them members of the council of the Professors' Association; 150 students were also expelled. The dismissal of the professors was a flagrant violation of the tenure system instituted in recent years.

Criticism has mounted to the point where Mr Federici was replaced by a new rector, Professor Juan de Dios Vial, on 29 October. We hope that this encouraging development will give rise to a reversal of the repressive measures instituted by the government. Meanwhile, we feel that the

international scientific community may be able to help our Chilean colleagues in their efforts to defend academic freedom and scientific research in Chile. With this purpose in mind, we ask the members of the scientific community to make known their opposition to the university closure and the abusive dismissal of the professors of the University of Chile by sending letters to the following: Professor Juan de Dios Vial, Rector, Universidad de Chile, Avenida Libertador Bernardo O'Higgins 1371, Santiago, Chile and Mr Juan Antonio Guzman, Ministro de Educación Pública, at the same address.

ALEXANDER MAURO

*The Rockefeller University,
1230 York Avenue,
New York, New York 10021, USA*

HERSCH M. GERSCHENFELD

*Laboratoire de Neurobiologie,
École Normale Supérieure,
46 Rue d'Ulm,
75230 Paris Cedex 05, France*

Refereeing reforms

SIR—The council of the American Physical Society (APS) has recently issued a statement, "Integrity of physics" (*Physics Today* 40, 81; 1987) which lists the following six actions as "sinful": (1) plagiarism, (2) data fabrication and manipulation, (3) submission of the same paper or its trivial variations to more than one publication, (4) fictitious coauthorship, (5) a reviewer's lack of impartiality and (6) slow response of a referee in order to suppress the publication.

There can hardly be doubts about the correctness of Commandments 1, 2, 4, 5 and 6, but I wonder about the third. Does it occur to the authors that this is in some contradiction with commandments 5 and 6? Under the existing refereeing system, the author of any innovative idea or experiment remains a helpless hostage of the anonymous referee(s) for an unspecified period of time. If, after some months, a paper is rejected (often almost without comment) the author has no means of claiming his priority — submission to another journal is a brand-new deal with, of course, a new submission date.

Simultaneous publication of exactly the same paper in two or more different journals is an embarrassing occurrence for the authors. Nobody wants this. Yet most people would agree that the simultaneous submission of 'trivial variations' of the same study to more than one journal is often the only practical way to reduce the risk of being victimized by the unfair refereeing system. This is especially true for a novice trying to enter a highly competitive area.

The solution which, I hope, many will find reasonable is the following. Major journals should promptly publish authors' abstracts of all submitted papers (unless

authors themselves instruct otherwise), leaving the acceptance of the full text subject to the usual refereeing. Length constraints and, possibly, an optional small charge could be applied. It would be for authors to ensure that their priority claims are properly expressed in the available space. I stress that the proposed system is not the same as the publication of abstracts in the *APS Bulletin*, for journals would publish only abstracts accompanied by full texts.

ALEXANDER A. BEREZIN

*Department of Engineering Physics,
McMaster University,
Hamilton, Ontario,
Canada, L8S 4M1*

Vatican and IVF

SIR — In reaction to Henri Firket's letter (*Nature* 329, 100; 1987) on the Vatican condemnation of *in vitro* fertilization, I would like to make the following points.

Firket defends the elimination of surplus fertilized eggs (a by-product of *in vitro* fertilization) on the ground that this is a common process in nature. Let us, however, suppose that the Vatican is right in its assertion that a fertilized human ovum is a human being in the normal sense of the word, although in its earliest stage of development. It becomes clear, then, that its "elimination" cannot be excused by pointing to nature's eliminating other human beings in later stages of their development, as is happening now, for example, in Ethiopia. Ethically speaking, humans sometimes have to act in a quite different way from what happens in nature.

Second, if this destruction of fertilized ova brought about by nature is necessary for the maintenance of the "genetic quality of the species", is maintaining this quality the aim of people who destroy

surplus fertilized ova?

Third, there is a passage I do not understand: "When this destruction of potential human life [the Vatican would say: leave out 'potential'] takes place as early as possible, it leads to a reduction of the total amount of suffering." Whose suffering? If the potential suffering of the unborn human being is meant, we are speaking about euthanasia in its earliest possible stage. If, however, the suffering of the survivors, who will experience the effects of population explosion, is meant, it does not explain why they have the right to live at the expense of other humans.

We must first decide whether a fertilized ovum is a human or not before we can talk about how to dispose of it. Until then, we should not dismiss the Vatican's view.

E. J. BONESCHANSCHER

*Vermeerstraat 18,
3817 De Amersfoort,
The Netherlands*

Genome sequencing

SIR — Further to David Tepfer's all too legitimate list of questions about the source of the sequenced human genome (*Nature* 329, 480; 1987), add the following. Should the person be male or female? Should he/she be alive or freshly dead (how much tissue do the molecular biologists need — an arm, a leg . . . ?) Should he/she be identified? Should candidates on the short list be screened for absence of heritable diseases?

And the most terrifying question of all, who is to decide?

MARTIN LUCK

*Institut für Hormon und
Fortpflanzungsforschung,
Grandweg 64,
2000 Hamburg 64, FRG*

A dim star with bright prospects

A single, almost invisible star can teach us a great deal about the Solar System, other stars and the rest of our Galaxy.

THERE are at least three reasons to be excited about the discovery, reported on page 138 of this issue by B. Zuckerman and E.E. Becklin, of a faint companion to the white-dwarf star Giclas 29–38. First, we do not know how many dim, low-mass stars there are in the Galaxy, so we cannot say what the total mass of the Galaxy is. Second, we do not know how much mass a self-gravitating body must accumulate before its internal pressure and temperature become great enough for nuclear fusion to begin and thus a star to be born. Third, we do not know what happens to the matter that is left over as a protostar condenses out of a gaseous cloud, and whether that matter is more likely to form another star, or planets, or a ring of debris like the asteroid belt. One discovery will not answer all these questions, but observers, knowing that stars exist even though we cannot see them, will be delighted that the range of visible stars has been extended into new territory.

A brown dwarf, which is what Zuckerman and Becklin say they have found, is an object whose mass is just too small for nuclear burning to begin in earnest (Liebert, J. & Probst, R.G. *A. Rev. astr. Astrophys.* **25**, 473–519; 1987). It never enters the main sequence of stellar evolution, and cannot properly be called a star. It will generate some internal heat both through gravitational collapse and by the burning of deuterium, but when that meagre supply of fuel is exhausted the object must slowly cool. During its 'hot' phase, a brown dwarf can reach a temperature of 2,000–3,000 K, and be visible as an infrared object. Zuckerman and Becklin's brown dwarf, which they saw as an infrared excess in the spectrum of an apparently normal white dwarf, seems to fit the bill fairly well. Its temperature is about 1,200 K, its luminosity 5×10^{-5} times the solar luminosity, and its mass 0.04–0.08 solar masses ($M_{\odot}$).

The uncertainty in the mass derives from uncertainties in the theoretical models with which the observations are compared. In any star, the internal temperature and density, and therefore the rate at which energy is generated, depend sensitively on the opacity of the stellar material, or in other words on the ease with which photons can escape from the core. In simple terms, the greater the opacity, the harder it is for heat to get out, making the core hotter. Opacities in normal stars can be calculated reliably

from a knowledge of the abundances of the constituent elements and their ionization states; but even so, different calculations of opacity in the Sun have significant differences in the predicted rate of solar neutrino emission. For brown dwarfs, the problem is harder because they are cool enough for simple molecules to exist, and finding the opacity is a problem involving chemistry as well as physics.

It is thought that a mass of about $0.01 M_{\odot}$ distinguishes planets, which do not burn, from brown dwarfs, and that a mass of $0.08 M_{\odot}$ separates brown dwarfs from true stars. Not only are these masses uncertain, but the boundaries are in any case fuzzy. An object near $0.08 M_{\odot}$ for example, can burn hydrogen slowly for a billion years, but then expire without ever reaching a true, nuclear-burning steady state. Careful scrutiny of Zuckerman and Becklin's brown dwarf could provide one good point on the stellar-evolution graphs.

Although brown dwarfs are invisible to us unless very close, they may contribute significantly to the total mass of the Galaxy. It has been recognized since the work of Oort in the 1960s that the dynamics of stars within a few parsecs of the Sun indicate the presence of more matter than can be accounted for by totting up all the visible stars. Very faint stars and brown dwarfs have always been popular candidates for this local 'missing' mass.

A histogram of the number of visible stars according to their mass shows that fainter stars, although individually less massive, are numerous enough to make up a larger proportion of the total galactic mass than bright stars. By continuing this trend, it is not difficult to suppose that a great deal of mass rests in brown dwarfs. On the other hand, the trend cannot continue to indefinitely small objects because they would harbour an infinite total mass. There must be a turnover in the distribution of stars at some low mass. The importance of finding this turnover has encouraged many searches for faint stars, with little success. An earlier detection by infrared speckle interferometry (McCarthy, D.W., Probst, R.G. & Low, F.J. *Astrophys. J.* **290**, L9–L13; 1985) of a brown-dwarf companion to the dim star Van Biesbrock 8 has, unfortunately, not stood the test of time: even its discoverers can no longer see it. Zuckerman and Becklin's discovery uses more conventional methods, and its greatest virtue may be to persuade others to keep looking.

There is no sound theoretical notion of how many low-mass objects, whether brown dwarfs or planets, we might expect to find. Because binary stars are so common, and because our own Solar System is so liberally provided with planets, moons, rings and asteroids, it is hard to escape the idea that star formation is a messy business, leaving lots of debris for other objects. But there is a more solid reason for supposing that stars are unlikely to form in isolation. As a cloud of gas begins to contract under its own gravity, any angular momentum it possesses will generate large circular velocities as the cloud shrinks, and unless the angular momentum can be disposed of, a disk, not a star, will form.

In the case of a binary star, such as Giclas 29–38, we can imagine that as the original gas cloud begins to turn into a disk, it solves its angular-momentum problem by splitting into two mutually orbiting bodies, which continue collapsing separately until they become stars. Most close binaries consist of stars of similar mass, which suggests a symmetrical formation process of this sort has occurred. On the other hand, if the total angular momentum of the collapsing cloud were smaller, it could condense further, into a flatter disk, before rotation put a stop to the process. At this point, magnetic fields could slowly transport angular momentum from the centre to the edge of the disk, allowing the central condensation to proceed towards star formation. The material in the outer disk can fragment and cool, so that planets, rather than a second star, will form. This fits the observation that the planets possess only one-thousandth of the mass, but nearly all the angular momentum, of the Solar System.

In this simple scheme, every star that is not part of a binary has planets around it, and indeed the general consensus today is that planetary systems are probably common. The low-mass binary discovered by Zuckerman and Becklin may be close to the borderline between the two cases. Spectroscopy may well reveal more brown dwarfs, whereas to find planets astrometry is needed; the careful tracking of stellar positions on the sky should reveal wobbles that betray the presence of an orbiting companion. But in either case, it will be good to see old-fashioned astronomical techniques, carried out on nearby stars, answering fundamental questions about the Galaxy, and the stars and planets in it.

David Lindley

Organelle transport

Retrograde step for microtubules

Jeremy S. Hyams

LOOKING at a living cell through the microscope is a bit like having a bird's eye view of Piccadilly Circus in rush hour. Vesicles and organelles show a complex two-way traffic pattern along intracellular highways of microtubules radiating out from the nucleus. Movement is either outwards (anterograde), towards the suburbs at the cell periphery, or inwards (retrograde), towards the bright lights at the cell centre. Once the black box of cell motility, organelle transport is rapidly becoming its star performer, at least as judged by the frequency of News and Views articles devoted to it. What has brought about this dramatic turn-around has been the ability to reconstitute particle movements *in vitro*^{1,2}. The first significant success of this approach was the identification of a protein, kinesin, which has at least some of the properties expected of the anterograde organelle translocator³. Now, in the September issue of the *Journal of Cell Biology*⁴ and on page 181 of this issue⁵, Richard Vallee and colleagues report the isolation of a protein with impressive credentials for being the retrograde organelle motor.

What is totally unexpected about their discovery is that this latest microtubule motor activity is not associated with a 'new' protein. Instead it is a hitherto unsuspected property of MAP 1C, one of the five high-molecular-mass microtubule-associated proteins or MAPs previously described by that laboratory⁶. These proteins (the others are MAP 1A, MAP 1B, MAP 2A and MAP 2B) were originally identified in brain. Their abundance varies markedly between different regions of the brain and from one tissue type to another, although the significance of this is still not fully understood. MAP 1C in particular has remained enigmatic because of its low abundance under all preparative conditions.

What Vallee and his co-workers now show^{4,5} is that, like kinesin, MAP 1C is greatly enriched in microtubules prepared in the absence of nucleotides. After selective removal of kinesin from their microtubule preparations using GTP, the authors could obtain a preparation highly enriched in MAP 1C by subsequent extraction of the microtubules with ATP. A microtubule-activated ATPase activity copurifies with MAP 1C, a clear indication that this protein might have a mechanochemical function⁴. To confirm this prediction, the authors turned to the remarkable ability of the AVEC microscope⁷ to see the movement of single microtubules. When microtubules are

placed on a glass microscope slide coated with MAP 1C, microtubule gliding occurs in a continuous, unidirectional fashion⁴. The rate of about $1.25 \mu\text{m s}^{-1}$ is comparable to that recorded for organelle motility in living cells, but about four- to fivefold higher than previously reported for kinesin using the same assay⁸.

Most significantly, however, MAP 1C and kinesin generate force in opposite directions⁵. The touchstone of microtubule-polarity experiments are the axonemes of the green alga *Chlamydomonas* in which the plus and minus microtubule ends can be unambiguously distinguished (see Fig. 3 of ref. 5, on page 183). On MAP 1C-coated slides, Vallee and collaborators found axoneme gliding is always from the compact (proximal or minus) to the frayed (distal or plus) end. Kinesin, on the other hand, causes axonemes to glide in the opposite direction. Because in living cells microtubules are oriented with their plus ends distal to the organizing centre at the nucleus, force production by MAP 1C is in the retrograde direction. Thus, MAP 1C could be the molecular motor for endocytosis, for fast retrograde transport in neurons and even for the positioning of organelles such as the Golgi apparatus and lysosomes near

the nucleus of cells.

Exciting as these findings are, they may be only the start of a much more profound insight into the mechanisms of intracellular transport and, indeed, of other forms of cell motility. Ever since the isolation of flagellar dynein almost 25 years ago and the demonstration of its role in force generation in ciliary and flagellar movement, there has been an intensive search for 'cytoplasmic dyneins', high-molecular-mass, microtubule-associated ATPases that might be expected to underlie the movement not only of organelles but also of chromosomes at mitosis and meiosis. MAP 1C may be just such a protein. Certainly, in terms of molecular mass, sedimentation coefficient, sensitivity to inhibitors and the finding that it can be specifically fragmented by ultraviolet light in the presence of vanadate, MAP 1C closely resembles flagellar dynein. Doubtless, experiments aimed at further establishing this relationship are already well in hand. Cell motility is once again a fast-moving field! □

1. Brady, T., Lasek, R.J. & Allen, R.D. *Science* **218**, 1129–1131 (1982).
2. Vale, R.D., Schnapp, B.J., Reese, T.S. & Sheetz, M.P. *Cell* **40**, 559–569 (1985).
3. Vale, R.D. *et al.* *Cell* **43**, 623–632 (1985).
4. Paschal, B.M., Shpetner, H.S. & Vallee, R.B. *J. Cell Biol.* **105**, 1273–1282 (1987).
5. Paschal, B.M. & Vallee, R.B. *Nature* **330**, 181–183 (1987).
6. Bloom, G.S., Schoenfeld, T.A. & Vallee, R.B. *J. Cell Biol.* **98**, 320–330 (1984).
7. Allen, R.D. *et al.* *J. Cell Biol.* **100**, 1736–1752 (1985).
8. Vale, R.D., Reese, T.S. & Sheetz, M.P. *Cell* **42**, 39–50 (1985).

Jeremy S. Hyams is Reader in Cell Biology in the Department of Biology, University College, London WC1E 6BT, UK.

Immunology

Synthetic peptides as vaccines

Jonathan Rothbard

THE concentration of effort to develop safe, cheap and effective vaccines to viral, bacterial and parasitic pathogens refocuses attention on several well-established tenets of immunology: the generation of the humoral response requires cooperation between B and T lymphocytes¹ (see figure); small-sized antigens induce immunity only when covalently attached to proteins that can be recognized by T cells^{2,3}; and immunological responsiveness is dependent on the HLA genotype of an individual⁴, which has now been correlated with the ability of an individual's histocompatibility (MHC) antigens to bind fragments of the immunogen⁵. Several recent papers describe the application of these principles to improve the use of synthetic peptides as vaccine components^{6–9}, the latest of which is reported on page 168 of this issue⁹ by Francis *et al.* of Wellcome laboratories. Whether these strategies will generate viable peptide vaccines remains to be seen, but the

experiments undoubtedly will generate useful research applications.

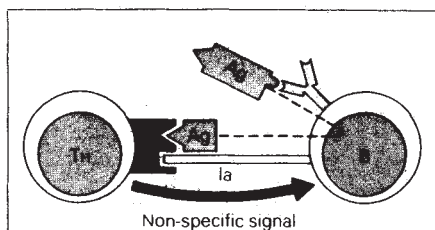
The groups attempting to develop vaccines against malaria⁶ and against foot-and-mouth virus⁷ have independently recognized that the linear peptides corresponding to protective antibody determinants are insufficient to immunize mice of divergent histocompatibility types. In either disease system, most strains of mice are non-responsive to the short linear peptides, but this problem can be overcome by covalently linking the peptide to a T-cell determinant capable of being recognized by the particular genetic strain of mouse.

In the case of malaria, Good and colleagues⁶ used a determinant within the circumsporozoite protein, whereas the Wellcome group⁹ conjugated a defined helper determinant from a heterologous protein to improve the immunogenicity of their foot-and-mouth peptide. Although these experiments are natural sequelae

from the work by Geftter, Gray and co-workers⁵, they clearly demonstrate that the importance of a carrier protein for a hapten is its ability to stimulate a helper T-cell response. The relative molecular mass of the carrier is unimportant, other than the larger the protein, the more likely it is to contain helper determinants for a greater number of histocompatibility types.

Perhaps a more important issue that these experiments collectively highlight is the problem that MHC restriction of the immune response poses for peptide vaccines. For although the non-responsiveness can be overcome for individual strains of mice, it is clearly a greater problem when considering the outbred human population. This concern also extends to proposed vaccines composed of individual viral proteins, because any one sequence is not guaranteed to contain regions recognized by all histocompatibility types.

One potential solution would be to



Possible mechanism for T-B cooperation. T-helper (T_H) cells are primed to antigen (Ag) associated with an MHC molecule (Ia) on antigen-presenting cells. The primed T cells then help B cells. Cooperation is restricted to the I region because of a requirement for the same type of MHC either on the T and B cells or on the T cell and the antigen-presenting cell that primes the T cell. (From ref. 14.)

generate a cocktail of peptides that can bind to most MHC class II alleles. Proponents of such a strategy can argue that such a cocktail would not necessarily be prohibitively large, both because the diversity of antigen-combining sites of class II proteins is limited and because certain sequences could be recognized by various histocompatibility types. Furthermore, most vaccines protect on the level of the population by reducing the rate of transmission and do not rely on protecting each individual. Consequently, a cocktail of peptides that would be recognized by a significant proportion of the population might be quite beneficial. But if I were going to an area where malaria was endemic, I would want to be vaccinated with a mixture that assured protection on the level of the individual.

Another possible solution to the problem of MHC restriction of the immune response was proposed recently¹⁰ by Milich and colleagues, who demonstrated that an antibody response to the nucleocapsid protein of hepatitis virus can be

generated in mice primed with synthetic peptides corresponding to helper T-cell determinants. Although there have been many unsuccessful attempts to generate similar stimulation using peptides corresponding to areas recognized by B cells, the success of Milich and colleagues is not surprising because T cells naturally recognize peptide antigens, whereas B cells bind proteins with their conformation intact. What is noteworthy in the experiments of Milich *et al.* is that by priming with a single T-cell determinant from the nucleocapsid, and subsequently boosting with the intact virus, the authors could generate antibodies against a second protein comprising the viral envelope.

Similar examples of such aberrant help have been seen previously in two other systems^{11,12}. The proposed mechanism does not violate cognate recognition of protein antigens by B and T cells. Rather, it postulates that immunogens that are aggregates, such as viruses, are treated by the immune system as single molecules. The polymeric structure is bound and internalized as a unit by B cells specific for any of the sterically available proteins of the virus, in this case, both the nucleocapsid and the envelope. Consequently, T-helper determinants from each of the viral proteins will be displayed on the surface of each B cell in association with MHC class II antigens (see figure). Helper T cells specific for the envelope protein generated from the priming with peptide then bind to B cells specific for each of the viral proteins, not just the envelope protein, resulting in the production of antibodies against both proteins. The aggregation of the two proteins is essential, shown by experiments in which primed mice generate antibodies only to the nucleocapsid when boosted with an equivalent mixture of the individual proteins.

The relevance of this work¹⁰⁻¹² to MHC restriction of the immune response is the demonstration that by priming a non-responding strain of mouse with a helper determinant from the nucleocapsid, antibodies against the envelope protein that the mouse normally cannot recognize are generated. Thus, priming with T-cell epitopes could be generally useful not only because the presence of pathogen-specific helper T cells will increase the kinetics of the immune response during an infection, but also because such a procedure might allow non-responsiveness to particular proteins to be circumvented by providing an alternative source of T-cell help.

In their latest work¹³, the Wellcome group has examined the usefulness of protein aggregates in immunization. The group greatly improved the immunogenicity of the foot-and-mouth peptide discussed above by fusing it to the hepatitis B nucleocapsid protein. This complex forms virus-like aggregates with the peptide

100 years ago

The scientific world, in the Duke of Argyll's opinion, has been for some time bowing down to the idol of Darwin and the theory of evolution, which is the fundamental dogma of that cult. Like a prophet of old the Duke raises a warning voice, and points out that the feet of the golden image are in part composed of clay.

Among the results of Mr Darwin's labours during the voyage of the *Beagle* in 1831-36 was a theory of the formation of Coral Reefs and Atolls. These are the Duke's words (*Nineteenth Century*, September 1887, p.305):

"Mr Murray's new explanation of the structure and origin of coral reefs and islands was communicated to the Royal Society of Edinburgh in 1880, and supported with such a weight of fact and such a close texture of reasoning that no serious reply has ever been attempted. . . The overthrow of Darwin's speculation is only beginning to be known. Can it be possible that Darwin was wrong? Reluctantly, almost sulkily, and with a grudging silence so far as public discussion is concerned, the ugly possibility has been contemplated as too disagreeable to be much talked about."

Prof. Huxley asserts that Darwin's confidence in the accuracy of his own theory was not seriously shaken, as the Duke alleges, and quotes as conclusive evidence a letter from Prof. Judd, who gives the results of a conversation which he had with Darwin no long time before the death of the latter. Prof. Huxley also intimates that Prof. Dana, "an authority of the first rank on such subjects," has pronounced against the new hypothesis.

So the "great lesson" has been read, and the scientific world, I fear, has not repented or rent its clothes. But it has heard, and not without indignation. The Duke of Argyll has made grave charges against the honour and good faith of men of science, and they ought to be grateful to Prof. Huxley for his prompt repulse of the attack and his stern rebuke of the assailant.

From *Nature* 37, 25; 10 November 1887.

displayed on its surface, and such aggregates are significantly more potent than when the peptide is administered in liposomes, emulsified in Freund's adjuvant or when fused to β -galactosidase. The authors do not know why their complex is more immunogenic than the other forms of the antigen, but the ability of hepatitis nucleocapsid to elicit a strong T-cell response distinguishes it from the other forms of polymerization. □

1. Claman, H., Chaperon, E. & Triplett, R. *Soc. exp. Biol. Med.* 122, 1167-1171 (1966).
2. Landsteiner, K. *The Specificity of Immunological Reactions* (Harvard University Press, 1946).
3. Mitchison, N.A. *Eur. J. Immun.* 1, 10-19 (1971).
4. Benacerraf, B. & McDevitt, H. *Science* 175, 273-279 (1972).
5. Guillet, S. *et al. Science* 235, 865-870 (1987).
6. Good, M. *et al. Science* 235, 1059-1062 (1987).
7. Borras-Cuesta, F., Petit-Camurand, A. & Fedon, Y. *Eur. J. Immun.* 17, 1213-1215 (1987).
8. Leclerc, C., Przewlocki, G., Chutsky, M. & Chedid, L. *Eur. J. Immun.* 17, 269-273 (1987).
9. Francis, M.J. *et al. Nature* 330, 168-170 (1987).
10. Milich, D.J., McLachlan, A., Thornton, G. & Hughes, J. *Nature* 329, 547-549 (1987).
11. Russell, S. & Liew, F. *Nature* 147, 280-282 (1979).
12. Lake, P. & Mitchison, N.A. *Cold Spring Harbor Symp. quant. Biol.* 41, 586-589 (1976).
13. Clarke, B. *et al. Nature* (in the press).
14. Roitt, I., Brostoff, J. & Male, D. *Immunology* (Gower, London, 1985).

Jonathan Rothbard is at the Imperial Cancer Research Fund Laboratories, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK.

Mass extinctions

Acid rain at the K/T boundary

Paul J. Crutzen

LIFE on Earth is in a constant state of flux, with new species emerging and others becoming extinct as a result of changing environmental pressures. The average life span of a species is 1–10 million years (Myr). The process is not steady, however, the most celebrated jolt being the Cretaceous/Tertiary (K/T) mass extinction 65 Myr ago. The popular explanation¹ of this event is that a massive meteorite collided with Earth and created a global dust cloud to block out the sunlight. Prinn and Fegley² now present a variant of this hypothesis: they suggest that a meteoritic impact made the global atmosphere 1,000 times more polluted than the worst modern smog, and that the precipitation of nitric acid caused the extinction.

The extinction pattern at the K/T boundary was complex and dramatic. The most famous victims of the K/T catastrophe were the dinosaurs, but 60–75 per cent of all marine species perished as well. Tropical species were the most strongly affected. Calcareous marine organisms were more vulnerable than siliceous ones. On land, reptiles and other animals heavier than 25 kg were hit the hardest. On the other hand, the inhabitants of lakes as well as terrestrial higher plants and many mammals did quite well. Clearly, any satisfactory model of the extinction event must not only explain the mass extinctions, but also the survival of many species (see refs 3–5).

Former hypotheses

Extraterrestrial causes for mass extinctions were not, until recently, considered likely by geologists. Older hypotheses included the effects of supernova events or huge solar flares during polar magnetic reversal periods. These hypotheses are relevant for the present discussion because they contain some important ingredients of mechanisms recently proposed. For instance, they involved large quantities of atmospheric nitrogen oxides, produced by bursts of X-rays, and galactic or solar cosmic rays. This would lead to two potentially catastrophic effects for the biosphere. First, solar radiation would be strongly absorbed by NO₂ in the atmosphere, decreasing its intensity at the Earth's surface, reducing drastically photosynthesis rates and surface temperatures. Second, nitrogen oxides (NO and NO₂) are efficient catalysts in the breakdown of stratospheric ozone, allowing harmful solar ultraviolet radiation to reach the Earth's surface. The problem is, however, that there is nothing in the geological or astronomical records to

substantiate such mechanisms.

Chemical analysis of the clay bed that separates the Cretaceous and Tertiary sediments showed that it is strongly enriched with iridium¹. L.W. Alvarez *et al.*¹ proposed that the iridium anomaly was caused by the impact of a huge meteorite (asteroid or comet) on Earth. Because the iridium was deposited world-wide, the meteorite must have been of the order of 10 km across. The energy of the impact, 10²³–10²⁴ J (equivalent to the solar radiation impinging on Earth over several weeks), must have raised atmospheric and surface temperatures by many degrees over a large part of the Earth. The impact was large enough to have lifted such enormous amounts of dust into the atmosphere that sunlight was blocked from reaching the Earth's surface. Alvarez *et al.*¹ also estimated that the mean time between such impacts would be about 30 Myr, remarkably close to the average period between mass extinctions determined by Raup and Sepkoski⁶.

The recent study by Prinn and Fegley², which follows earlier ideas of Prinn and colleagues⁷, addresses the question of the selectivity of the extinctions at the K/T boundary. They suggest that a very large cometary impact could produce up to 10¹¹ molecules (3 × 10¹⁸ g) of NO (a factor of ten less seems to me more reasonable, based on the estimated impact energy). This is produced not so much by the bolide itself, but by the shock waves created by the high-velocity plume of terrestrial debris that would be ejected from the ocean and the submarine crust. The authors estimate that at least one-eighth of the original bolide energy is transferred to the atmosphere, mostly by small particles. After global spread through the atmosphere, which should take less than a year, the entire atmosphere is loaded with 100 parts per million by volume (p.p.m.v.) of NO₂, about 1,000 times more than during the heaviest air-pollution episodes in some of our modern cities. The result would be total darkness at the Earth's surface, global climatic cooling and severe disruptions of the food chain.

This, however, would not be the only catastrophic environmental effect of such an impact. Prinn and Fegley point out that 100 p.p.m.v. of NO₂ is enough to poison any animals and plants directly exposed to the atmosphere. Furthermore, the nitrogen oxides would react with water during the following months, to produce large amounts of nitric acid, causing superacid rainfall (with a global-average pH of about 1), destroying much of the

biosphere. In particular, the authors state that the precipitation of the nitric acid would lower the alkalinity of the top 75 m of the ocean so much that many calcareous oceanic organisms would dissolve. Although this acidity could be partially neutralized by the plume of debris, most of the formation and deposition of HNO₃ would occur after the particulates had settled on the ocean surface. Organisms living in shallow water, well buffered by carbonate rocks against acidification, are well protected against extinction. The most sensitive faunal species are terrestrial vegetarians and animals laying eggs exposed to the atmosphere. Species that could hibernate or lay eggs in burrows are much better off, which is consistent with the geological records. Prinn and Fegley's model, therefore, could explain the preferential extinction of calcareous surface species, and the survival of many deep-sea and siliceous organisms.

Recovery

Although extensive portions of the world's forests would be destroyed by the heat and shock waves generated by a large impact and by the severe meteorological chemical and climatic conditions that would follow, many species could re-establish themselves from seed banks. (If world-wide forest fires took place, as can be inferred from the deposition of about 10¹⁷ g of soot in the K/T boundary layer⁸, fire-resistant species may have been strongly favoured.) Prinn and Fegley further point out that HNO₃ deposition on carbonate-poor regions can mobilize large quantities of highly toxic trace metals in the environment, especially in coastal waters. In this context, they suggest that their extinction mechanism can be tested by measuring the trace-metal content of fossils of neritic organisms in the K/T boundary clays.

Another important environmental factor that could influence the re-establishment and evolution of the biosphere is the atmospheric carbon dioxide level. With the oceanic carbon cycle severely perturbed by the removal of many organisms from the ocean surface, the concentration of atmospheric carbon dioxide could have been substantially elevated over millenia, making possible extended greenhouse warming. Note that Emiliani *et al.*³ propose that most extinctions occurred because of a global heat wave for which tropical species were ill prepared. The cause of the heat wave was supposedly the 'greenhouse' effect resulting from the injection of large amounts of H₂O into the atmosphere.

Prinn and Fegley do not address the fate of the huge quantities of fixed nitrogen added to the Earth's surface by such an event, equivalent to many thousands of years of biological nitrogen fixation. How much of it could be denitrified to produce

N_2O and NO , and how quickly? If NO were produced, and escaped to the atmosphere, the effects indicated by Prinn and Fegley could last substantially beyond the few years that they suggest. Furthermore, substantial ozone could be produced in the troposphere by smog reactions, involving the oxidation of CH_4 produced by the large quantities of anaerobically decaying organic matter.

Greenhouse gases

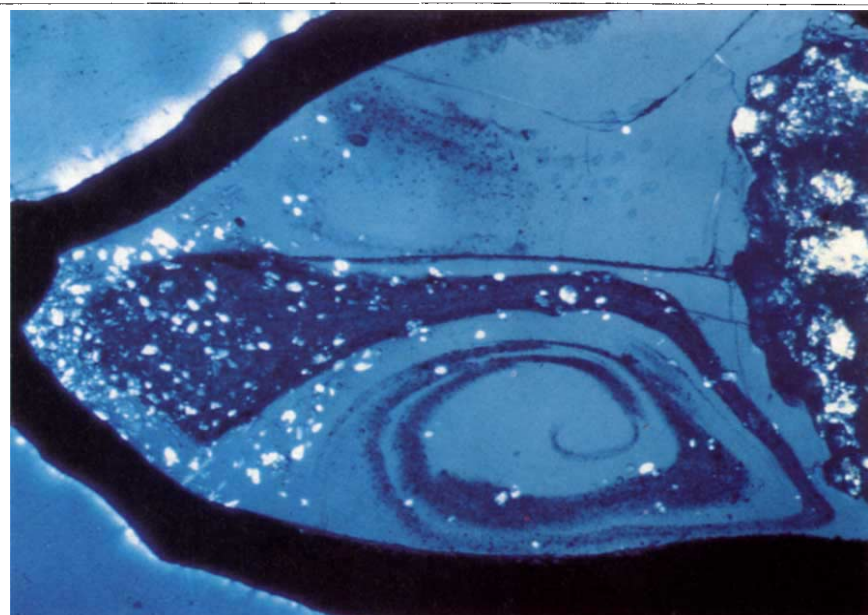
If N_2O were produced, the effects would be very different. The lifetime of N_2O in the atmosphere is of the order of 150 years, and its dissolution in the oceans is much lower than that of CO_2 . Could the post-impact atmosphere contain high concentrations of N_2O together with CO_2 , HNO_3 and CH_4 ? If so, the Earth could warm up considerably, because these are important greenhouse gases. At the same time, stratospheric ozone could be substantially reduced, as the oxidation of N_2O is the main source of the ozone-destroying catalysts NO and NO_2 .

It is important to consider the implications of the discovery of 10^{17} g of soot in the K/T boundary clays⁸. Extrapolating from known emission factors of forest fires, more than 10^{19} g CO_2 , 10^{18} g CO , 10^{17} g CH_4 , reactive hydrocarbons and oxides of nitrogen ($\text{NO} + \text{NO}_2$) and 10^{16} g N_2O should have been produced as well. This would have led to there being about 15 times more CO_2 , 4,000 times more CO , 30 times more CH_4 and 7 times more N_2O in the atmosphere than at present. The result would be an enormous greenhouse warming (surface temperatures raised by about 10°C) and an intense world-wide photochemical smog. It is, therefore, of critical importance to identify the source of the soot in the K/T boundary layer (forests or fossil fuels) and especially over what length of time the soot production took place.

The atmospheric chemical conditions during the centuries or millenia after the meteoritic impact may well have been at least as important for the further evolution of life as the shorter-term perturbations caused by the impact itself. Many species could not survive. Those that did must have enjoyed a paradise. □

1. Alvarez, L.W., Alvarez, W., Asaro, F. & Michel, H.V. *Science*, **208**, 1095–1108 (1980).
2. Prinn, R.G. & Fegley, B. Jr *Earth planet. Sci. Lett.* **83**, 1–15 (1987).
3. Emiliani, C., Kraus, E.B. & Shoemaker, E.M. *Earth planet. Sci. Lett.* **55**, 317–334 (1981).
4. Raup, D.M. & Jablonski, D. (eds) *Patterns and Processes in the History of Life* (Springer, Heidelberg, 1986).
5. Elliot, D.K. (ed.) *Dynamics of Extinction* (Wiley, New York, 1986).
6. Raup, D.M. & Sepkoski, J.J. Jr *Science* **231**, 833–836 (1986).
7. Lewis, J., Watkins, H., Hartmann, H. & Prinn, R.G. *Geol. Soc. Am. Spec. Pap.* **190**, 215–221 (1982).
8. Wolbach, W.S., Lewis, R.S. & Anders, E. *Science* **230**, 167–170 (1985).

Paul J. Crutzen is at the Max Planck Institute for Chemistry, PO Box 3060, D6500 Mainz, FRG.



Most natural granite melts are too viscous to flow except on the largest scale, yet millimetre-scale flow has been observed in molten granites in a recent experiment. Michael R. Carroll and Peter J. Wyllie at the California Institute of Technology performed an experiment intended to observe the contamination of the granite melt (light blue) by peridotite (dark blue, far left) at conditions typical of the crust–mantle boundary. Convection was not expected, and is visible only because of tiny grains of graphite formed by the reduction of CO_2 present in the starting materials. Crystals of quartz and feldspar (white) also became entrained in the flow. The viscosity of the sample was reduced by the presence of water in the starting materials, but was still too high for the flow to have been driven by thermal convection, because the vertical temperature gradient was too small. Other mechanisms involving volume changes caused by crystallization or dissolution, or variations in composition or surface tension, may have been at work. (Sample $2\text{mm} \times 1\text{mm}$). □

Cosmic rays

Modulation in three dimensions

J.R. Jokipii

FLUCTUATIONS in the flux of galactic cosmic rays were first attributed to solar effects 40 years ago. The strongest variation occurs in anti-phase to the 11-year sunspot cycle. Observations reported at a recent conference* now show that the latitudinal gradient of cosmic rays is correlated with the large-scale magnetic structure of the heliosphere.

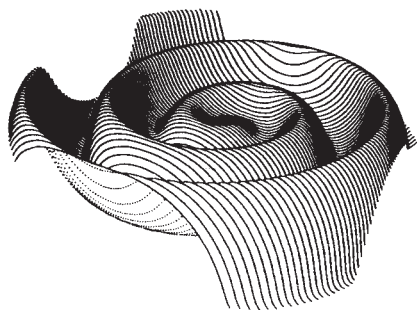
The solar modulation was first attributed by Parker (*Phys. Rev.* **110**, 1445; 1958) to the solar wind and the magnetic field trapped within it. The solar wind, blowing radially at several hundred km s^{-1} , creates a cavity, the heliosphere, in the interstellar gas. In the process, it drags out the solar magnetic field in the form of a spiral, resulting from solar rotation. Parker's original idea was that magnetic irregularities, carried in the wind, cool the cosmic rays and exclude them from the inner Solar System. The observed cosmic-ray flux results from a balance between inward diffusion, cooling and convection. This simple picture neglects the coherent drift of the particles, which occurs if the field has a large-scale coherent structure.

This structure was observed by Pioneer-11 as it moved to 16° heliographic latitude (Thomas, B.T. & Smith E.J. *J. geophys. Res.* **86**, 11,105–11,110; 1981).

For several years around each sunspot minimum, the magnetic field is organized simply. Around the 1975 minimum, the field in the Northern Hemisphere spiralled out from the Sun, and that in the south spiralled inward. The two hemispheres were separated by a thin current sheet near the heliographic equator. The magnetic fields alternate in sign at successive sunspot minima and at present the northern field is directed inwards. Thus, at least near sunspot minima, the magnetic field has a large-scale coherent structure, and drifts may be important. The situation is still not clear near sunspot maxima.

Theoretical research and numerical models over the past decade have explored the consequences of the gradient and curvature drifts of cosmic rays in the newly established large-scale interplanetary magnetic field. A firm prediction of theory, based on reasonable extrapolations of ecliptic observations to higher latitudes, is that latitudinal drift changes things fundamentally. In fact, theory

*20th International Cosmic Ray Conference, Moscow, 2–15 August 1987.



Simulated image of the undulating heliospheric current sheet as seen by an observer 30° above the equatorial plane and 75 AU (11×10^9 km) from the Sun. The figure is 25 AU across. (From Jokipii, J.R. & Thomas, B.T. *Astrophys. J.* **243**, 1115; 1981.)

predicts that, when the northern magnetic field spirals outwards (around 1975), positively charged particles reaching the inner Solar System propagate inwards and latitudinally from the outer, polar regions of the heliosphere, whereas now (northern field inwards) they are expected to propagate inwards from the equatorial regions, along the current sheet. This would result in a 22-year intensity cycle for cosmic rays as well as effects that depend on the charge of particles. Neither was predicted by the earlier models. In addition, a basic prediction of drift theory is that the particle intensity should increase with heliographic latitude during the 1975 solar minimum and decrease during 1985. These controversial predictions do provide a natural interpretation of important observations (see my article in *The Sun and Heliosphere in Three Dimensions* (ed. Marsden, R.G.) 375–387; Reidel, Dordrecht, 1986).

Alternative theories, which rely solely on radial transport, invoking various magnetic barriers to the cosmic-ray motion, remain popular in some quarters. There is little doubt that some of these effects occur and cause at least local changes in the cosmic-ray flux. But there is as yet no fully comprehensive, consistent global theory of heliospheric transport.

The unravelling of this story is complicated by the fact that spacecraft are confined to near the ecliptic, as the most important new effects occur as a function of heliographic latitude. But the picture is improving. Voyager 1, at a radial distance of some 25 astronomical units (AU) has moved to nearly 30° heliographic latitude. Data from Voyager 1, Voyager 2 (near the equatorial plane), Pioneer 10 (at 40 AU radius) and the Interplanetary Monitoring Platform (IMP) and the International Sun-Earth Explorer (ISEE), both near Earth, provide an unprecedented opportunity to study the latitudinal structure. Also the observations were fortunately made near the time of a sunspot minimum, with the attendant simplification of the magnetic structure of the heliosphere.

Measurements of the latitudinal gradient of cosmic rays obtained from these

spacecraft were reported at the recent meeting in Moscow (E.C. Stone, Caltech; F.B. McDonald, NASA; D. Venkatesan, Calgary Univ.). There is a clear consensus that, at the present sunspot minimum, the intensity decreases away from the heliospheric current sheet (and from the equatorial plane of the Sun), much as predicted by theory. Moreover, there is evidence from previous Pioneer data (McKibben, R.B. *et al.* *Astrophys. J.* **227**, L147–L152; 1979), that the intensity increased away from the equatorial plane during the last sunspot minimum. This change in sign of the latitude gradient is a firm prediction of the drift theory, but not of the old theory without assumptions.

The drift theory still has difficulty in explaining some basic observations, including the temporal behaviour of cosmic-ray electrons and the lack of change in the radial dependence of the galactic cosmic-ray intensity between successive sunspot cycles. It could be that

simply changing the model parameters improves agreement with observations, or perhaps exotic new effects such as a net helicity of the interplanetary turbulence are involved. The poorly understood outer boundary of the heliosphere may also be important (J.A. Simpson, Chicago Univ.; M.A. Forman, Stony Brook).

Even with these new insights, direct observations are still restricted to what is effectively the equatorial zone of the heliosphere. Our present knowledge is comparable to an understanding of the Earth that would be obtained by people confined to within 30° of the equator. The Ulysses mission, to fly over the poles of the heliosphere during the next decade, will significantly increase our understanding of the three-dimensional nature of modulation and of the heliosphere. $\square$

J.R. Jokipii is in the Planetary Sciences and Astronomy Departments of the University of Arizona, Tucson, Arizona 85721, USA.

Synaptic physiology

Inhibition by acidification?

Roger Thomas

NERVE cells usually inhibit their targets by releasing a chemical transmitter that opens chloride channels in the post-synaptic cell. Because the chloride equilibrium potential is close to the resting potential, the open channels make it harder for excitatory inputs to stimulate the cell. At least some types of chloride channel are now known¹ to be permeable to bicarbonate as well as to chloride ions. Despite this, the possibility that bicarbonate fluxes through chloride channels might change intracellular pH (pH_i) has not hitherto been discussed, let alone investigated. This omission is now elegantly rectified by the report on page 163 of this issue² by Kaila and Voipio. The authors show that bicarbonate efflux through GABA-activated inhibitory channels can cause a significant fall in pH_i in crayfish muscle fibres. As GABA- and

glycine-activated channels in mammalian central neurons are also permeable to bicarbonate, these new findings have wide implications. They also suggest that older conclusions on chloride channel selectivity need to be re-examined.

The mechanism by which GABA (γ -aminobutyric acid) causes a decrease in pH_i in crayfish muscle is shown in Fig. 1. Bicarbonate ions leaving via the anion channel are replaced by the reaction, catalysed by carbonic anhydrase, of the omnipresent CO_2 with water. This makes more bicarbonate and also H^+ ions. Each exciting bicarbonate in effect generates an extra H^+ ion, which causes the decrease in pH_i . Kaila and Voipio also report an increase in the surface pH, presumably caused by the bicarbonate combining with surface H^+ ions to produce CO_2 . *In vivo*, the intracellular bicarbonate level is determined primarily by the pH_i and CO_2 level, and will normally be about 15–20 mM in both crayfish muscle and mammalian neurons^{3,4}, only slightly less than the extracellular level. Thus at the resting potential there is a large electrical gradient available to drive a bicarbonate efflux. When chloride channels are opened by GABA, bicarbonate ions will exit through them and chloride ions will tend to enter.

Another important consequence of the bicarbonate flux revealed by Kaila and Voipio's pH_i measurement is that the reversal potential for the GABA response is not the same as the chloride equilibrium potential, which is close to the resting potential. Thus, in the presence of

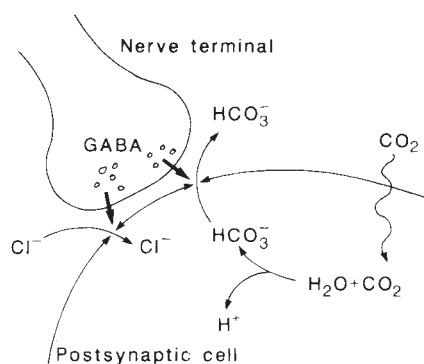


Fig. 1 How bicarbonate efflux through a GABA-activated anion channel will cause a fall in postsynaptic pH_i .

bicarbonate the inhibitory potentials reverse at potentials 10–15 mV more positive than the chloride equilibrium potential. In the past the reversal potential of the inhibitory postsynaptic potential⁵ has been used to estimate intracellular chloride levels. These estimates will almost certainly be too high if they are made in the presence of CO₂ and bicarbonate, as they would be *in vivo*.

Are all chloride channels likely to be permeable to bicarbonate? Early researchers may have been ignorant of the instability of injected bicarbonate ions. Ever since the pioneering days of Coombs and co-workers⁶, synaptic physiologists have investigated the nature of the inhibitory channel by injecting various anions into neurons and observing the effects on the potential at which the inhibitory response reverses. Anions that shift the reversal potential towards zero were thought to permeate the channel, whereas those with no effect are generally larger and therefore were presumed to be too big. Bicarbonate injection has no effect, suggesting that the channel has an effective diameter of about 3.4 Å. More recently, by a different injection method, Bormann *et al.*¹ have shown that bicarbonate can permeate both glycine- and GABA-activated chloride channels in

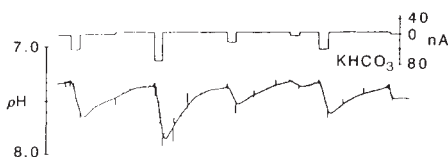


Fig. 2 The effect of five separate 2-min bicarbonate injections (upper trace shows injection current) on the pH of a snail neuron superfused with a bicarbonate-free saline. The bicarbonate takes up a H⁺ ion and leaves as CO₂, causing a rapid increase in pH_i. (Redrawn from ref. 8.)

mouse neurons. The estimated diameter for glycine-activated channels is 5.2 Å.

I have looked in vain for signs of awareness of the lability of bicarbonate, and indeed of other anions of weak acids such as acetate and propionate, when injected into nerve cells. Perhaps earlier conclusions (including my own⁷) that these ions have no effect on inhibitory potentials simply resulted from the fact that the ions vanish from the cell interior as fast as they are injected. Thus, all that may be achieved by the injection of these anions is to cause a rise in pH_i, as shown in Fig. 2. Even where these ions do change the inhibitory response, there must also be some increase in pH_i, with unpredictable effects on membrane properties.

The alternative explanation of the discrepancy between early and recent estimations of channel diameter is that anion channels of the cat motoneuron are smaller than those of the mouse spinal neuron. But because the same transmitter (glycine) is responsible for inhibiting cat

and mouse neurons, this explanation seems unlikely. Perhaps all inhibitory synaptic channels are permeable to bicarbonate as well as to chloride. If so, Kaila and Voipio's results suggest that all postsynaptic inhibitory ion fluxes will generate changes in intracellular and, possibly more significantly, in surface pH. It is often forgotten that bicarbonate is always present *in vivo*. Whether these pH changes are ever large enough to have any biophysical effect is unknown, but it is at least an intriguing possibility that hyper-

active synapses might inhibit or modulate by causing local pH changes. □

1. Bormann, J. *et al.* *J. Physiol., Lond* **385**, 243–286 (1987).
2. Kaila, K. & Voipio, J. *Nature* **330**, 163–165 (1987).
3. Roos, A. & Boron, W.F. *Physiol. Rev.* **61**, 296–434 (1981).
4. Thomas, R.C. *J. Physiol., Lond* **354**, 3–22P (1984).
5. Eccles, J.C. *The Physiology of Synapses* (Springer, Berlin, 1964).
6. Coombs, J.S. *et al.* *J. Physiol., Lond* **130**, 326–373 (1955).
7. Kerkut, G.A. & Thomas, R.C. *Comp. biochem. Physiol.* **11**, 199–213 (1964).
8. Thomas, R.C. *J. Physiol., Lond* **255**, 715–735 (1976).

Roger Thomas is Head of the Physiology Department, University of Bristol, Bristol BS8 1TD, UK.

Nuclear structure

Near-harmonic vibrations found

Kris Heyde

ATOMIC nuclei are typically discussed in terms of either the shell model or the liquid-drop model, depending on the number of protons and neutrons present. The shell model considers the motion of nucleons in an average spherical field with additional residual interactions. The liquid-drop model describes the collective motion of the nucleons. A benchmark of this model is the spherical-quadrupole-vibrator model of A. Bohr and B. Mottelson (*K. danske Vidensk. Selsk. Skr.* **24**, No. 16; 1953). Only now have the nearly harmonic vibrations, characteristic of this model, been observed conclusively in a real nucleus (Arahamian, A. *et al. Phys. Rev. Lett.* **59**, 535–538; 1987).

Nuclei that are described in the shell model as having only a few valence nucleons should best fit the spherical-quadrupole description, because of their spherical shape and 'soft' potentials. The oscillations on such a potential are nearly harmonic and the excitations are almost degenerate, depending mainly on the number of phonons of an excitation and only slightly on the angular momentum.

Although for many even-even nuclei (those with even numbers of protons and neutrons), the two-phonon excitation has almost twice the energy of the single-phonon excitation, one characteristic of harmonic oscillations, few show the triplet near-degeneracy expected in the two-phonon excitation. In some cases, some states are missing (perhaps for experimental reasons), or splitting of the triplet levels is too large and extra levels occur (such as non-collective excitations or collective rotation-like excitations). There has been no clear-cut evidence of higher phonon multiplets until now.

In their recent paper, Arahamian *et al.* show that the nuclide cadmium-118 (¹¹⁸Cd) has a nearly degenerate two-phonon triplet of levels, and probably a full quintuplet of three-phonon levels. The cadmium nuclei were studied by observing the radioactive decay of silver nuclei

produced at the TRISTAN isotope separator at Brookhaven National Laboratory by fission of ²³⁵U (Fig. 1).

The levels within the triplet are split by 120 keV at most, which is small compared with the excitation energy of approximately 1.2 MeV. The identification of the individual levels within the quintuplet are less certain (the angular momentum of two not being uniquely determined), but it is very likely that these are the five expected three-phonon excitations. Both the energies and the decay patterns of the levels correspond well to the harmonic-oscillator model of the nucleus.

The ¹¹⁸Cd isotope was chosen as a candidate harmonic vibrator after systematic studies of other even-mass cadmium isotopes (see Fig. 2). Spectra showed that additional low-lying states (labelled 0⁺ and 2⁺) disturb the two-phonon levels in isotopes with mass number less than 118. For the heavier isotopes, some additional levels have higher energies, falling between the two- and three-phonon levels, so that the perturbation is minimized.



Fig. 1 The isotope separator TRISTAN at Brookhaven National Laboratory. Radioactive silver nuclei (¹¹⁸Ag) produced by fission of ²³⁵U are ionized, extracted from the source, mass-selected and trapped on movable tape in the detection station (front, centre right). The cadmium spectrum was measured from the γ-emission following the β-decay of the silver nuclei. The angular momenta were deduced from angular correlations of emitted γ-rays.

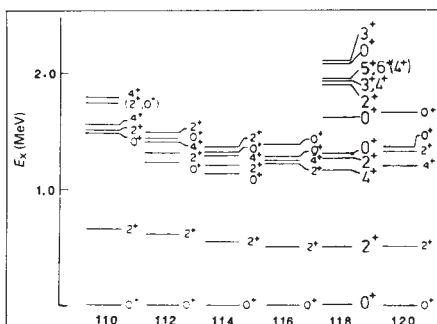


Fig. 2 The low-lying levels of cadmium isotopes ^{110}Cd – ^{120}Cd , as identified spectroscopically. E_x , excitation energy. The ground-state (0^+), one-phonon level (2^+) and triplet two-phonon levels (0^+ , 2^+ and 4^+) are shown, together with the additional 0^+ and 2^+ single-particle excitations close to the triplets. Closely lying states with the same symmetry become quantum-mechanically mixed, so that the additional levels take on some two-phonon character. The spectrum of ^{118}Cd also shows the five three-phonon levels (0^+ , 2^+ , 3^+ , 4^+ , 6^+) observed by Aprahamian *et al.* In spectroscopic notation, the number refers to the angular momentum of the nuclear state, and the superscript (+ or –) indicates its 'parity'.

The possibility of observing the three-phonon levels is somewhat surprising, as the excitation energy is close to the energy needed for the breaking up of nucleon pairs. For nuclei, such as cadmium, with approximately 50 protons (one of the 'magic' numbers of added stability in shell theory), the energy needed to break up pairs is about 2 MeV, twice the 'pairing-gap' energy. For cadmium, however, there being mainly one valence level involved (the $1g_{7/2}$ orbital), there are few such excitations possible at lower energy and these incoherent states are thus unlikely to perturb the coherent three-phonon levels. Other nuclei near a magic number, such as tellurium, have a more complicated single-particle spectrum, and the resultant complications easily destroy the underlying vibrational spectrum at the three-phonon level.

Aprahamian *et al.* point out that their spectrum can also be accounted for within the vibrational limit (or U(5) limit in the language of group theory) of the interacting boson model of F. Iachello and A. Arima (*The Interacting Boson Model*. Cambridge Monographs on Mathematical Physics, 1987). This limit allows for anharmonic effects in multi-phonon excitations and so describes a larger class of nuclei. Both the excitation energies and the splitting within multiplets are well reproduced, and there is some evidence of four-phonon excitations. Further work on these, as well as a full identification of the angular momentum of the three-phonon levels, should establish ^{118}Cd as a model liquid-drop-type nucleus. Perhaps a similar study of ^{120}Cd should follow. □

Kris Heyde is at the Nuclear Physics Laboratory, Gent State University, Proeftuinstraat 86, B-9000 Gent, Belgium.

Sir Peter Medawar (1915–1987)

PETER MEDAWAR, who died on 2 October, was the most distinguished British biologist of his generation. He founded transplantation biology as a scientific study, and thus laid a cornerstone of modern immunology. Under his influence, immunology became a respectable subject for biologists to enter.

His favourite experiment was to measure with meticulous precision the time taken for rejection of a skin graft. In this way he discovered that second-set grafts undergo accelerated rejection, and hence that rejection is brought about by an immunological response on the part of host lymphocytes against histocompatibility antigens in cells of the graft. This led on to identification of the lymphocyte as the immunologically reactive cell, and eventually to recognition



Sir Peter Medawar — distinguished discoveries.

that cellular immune responses are as important as antibodies to the immune system. Having identified the mechanism of rejection, he went on to discover, in the anterior chamber of the eye and elsewhere, sites where lymphocytes do not penetrate and which therefore enjoy 'immunological privilege'.

Medawar's greatest single discovery, made jointly with Rupert Billingham and Leslie Brent (and contemporaneously with independent work of Ray Owen's), was of immunological tolerance: the immune system is not pre-programmed to distinguish between self and non-self, but learns to do so as a result of exposure to self-molecules during early development. This

principle enables us to understand how the immune system is shaped, helps establish the working of clonal selection among lymphocytes, and leads on to a rational understanding of auto-immune disease. For this work, Medawar shared the Nobel prize for physiology or medicine in 1960. Thirty years later we are still arguing about just how much clonal deletion occurs in establishing a state of immunological tolerance.

The war effort and the needs of burn victims drew Medawar away from his first work, which was on the growth of embryos. He retained a strong interest in developmental biology, and his discovery with Billingham of pigment spread later persuaded him to follow the plasmagene will o'the wisp. What fascinated him was the possibility that pigmented cells might transmit into their non-pigmented neighbours self-replicating information for pigment production: alas, pigmentation turns out to spread only through cell movement.

As director of the National Institute for Medical Research between 1962 and 1971, his research focused on the development of anti-lymphocyte serum to prevent graft rejection — a sledgehammer used to crack a nut? After leaving the institute he returned finally to embryos and elegantly demonstrated cross-immunity between cancer cells and fetal tissue. He was devoted to the theory of anaplasia, which he claimed would have been widely accepted had not most cancer biologists believed her to be a defunct Imperial Russian princess.

Medawar and molecular biology is a subject that deserves an essay on its own. A good try, but a little too early? Let down by colleagues who should have made the connection (myself included)? Too difficult for an Oxford biologist of his generation, raised on genetics as a personal game played between Ford, Fisher and Haldane, to take the subject altogether seriously? Certainly Medawar warmly welcomed the advent of molecular genetics: witness his review of *The Double Helix*. But from the subject of transplantation immunology to its successor, immunogenetics, Medawar never personally quite managed to bridge the gap.

Peter Medawar was endowed with a sense of humour, a profound respect for the creative individual, and a dislike of pomposity. One has no difficulty in imagining what he would have made of current notions on "central initiatives" and "University Research Centres". His own research was invariably carried out within a small, personally devoted group.

N.A. Mitchison

N.A. Mitchison is in the Department of Zoology and Comparative Anatomy, University College London, London WC1E 6BT, UK.

Plasma lipoproteins

Teaching old dogmas new tricks

Michael S. Brown and Joseph L. Goldstein

PLASMA lipoproteins, *agents provocateurs* of atherosclerosis, are implicated in half of all deaths in industrialized countries. The penumbra surrounding these lipid-laden troublemakers is finally being stripped away by the techniques of molecular biology. The two latest revelations concern the largest and most enigmatic protein of the lipid-transport system, apoprotein B (apo B). One is described in recent issues of *Cell*¹ and *Science*²; the other is described on page 132 of this issue³ by the collaborative efforts of Anglo Scanu and Gunther Fless at the University of Chicago and a team led by Richard Lawn at Genentech. Both these discoveries challenge contemporary ideas of biology: the first challenges the dogma that mature messenger RNAs are inviolate; and the second challenges the dogma that evolution makes sense.

Lipoproteins are composed of a fatty core of cholesteryl esters and triglycerides, surrounded by a monolayer of phospholipid and cholesterol in which various proteins (called apoproteins) are embedded (see figure). The triglyceride-rich lipoproteins, chylomicrons and very low-density lipoproteins (VLDLs), are secreted by the intestine and liver, respectively. Both lipoproteins contain apo B, which is essential for assembly of the hydrophobic core.

Intestinal and hepatic apo B differ in size⁴. The hepatic form, found in VLDLs, is huge, containing 4,563 amino acids, and called apo B-100. The intestinal form, found in chylomicrons, contains only the amino-terminal 2,152 amino acids of apo B-100, and is called apo B-48 (it is 48 per cent as large as apo B-100)^{1,2,4}. In the bloodstream, VLDLs and chylomicrons are acted on by lipoprotein lipase, which removes the triglyceride. This reaction converts the VLDLs to LDLs, which are cleared from the circulation when the carboxy-terminal domain of apo B-100 binds to LDL receptors⁵. Apo B-48, which lacks the carboxy-terminal domain, does not bind to LDL receptors. In chylomicron remnants, receptor binding is mediated by another protein, apo E, together with apo B-48.

In a rare recessive disease, abetalipoproteinaemia, apo B-48 and apo B-100 are absent from blood⁴, suggesting that both proteins are produced from the same gene, perhaps through alternative splicing

of the messenger (m) RNA. But the cloned gene encoding apo B-100 shows no evidence for alternative splicing. Indeed, the position at which apo B-48 terminates lies in the middle of an enormous exon of 7,572 base pairs, the largest in any gene⁶.

In their recent papers, Powell *et al.*¹ and Chen *et al.*² report the solution to this enigma, with startling implications. They used reverse transcriptase to clone 22 apo B-48 complementary (c) DNAs from the intestines of three humans and one rabbit. In all these clones, the mRNA was

dependent on apo E, and this allows these dietary particles to be diverted to a receptor that is different from the receptor that clears LDL.

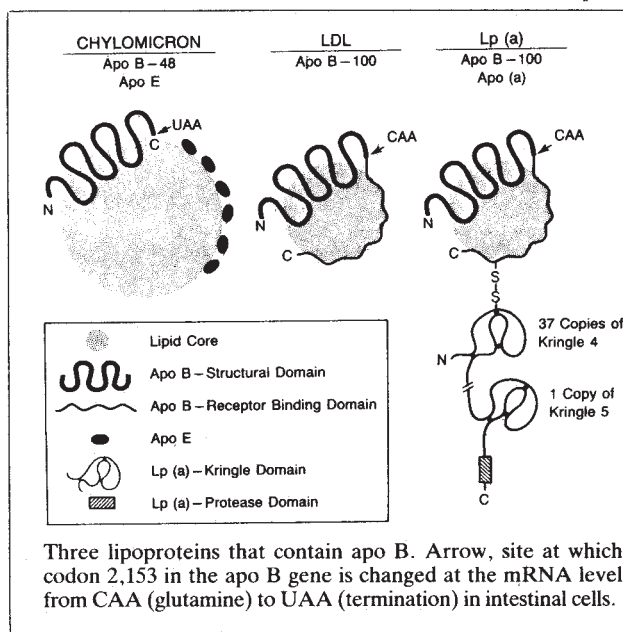
The central dogmas of molecular biology have been falling one by one. The discoveries of reverse transcriptase, introns, genomic rearrangements and catalytic RNA show that the transmission of genetic information is more plastic than previously thought. Yet despite these results, one structure has — until now — remained inviolate. Once an mRNA is formed, its coding region is supposed to remain intact, or else chaos might intervene. The discovery of a 'sense-nonsense' conversion in the apo B mRNA raises the possibility of a new type of mRNA editing. We must now learn whether this process exists only for one apoprotein gene, or whether nature has found other genes in which to put this mechanism to use.

The second structural revelation about apo B, described by McLean *et al.* on page 132 of this issue³, is equally astounding. It involves a variant of LDL called Lp(a), which was discovered in 1963 by Kaare Berg as an antigen in the blood of certain individuals. Lp(a) is an LDL that carries one copy (or perhaps two) of a protein called apo(a), joined to apo B-100 by a disulphide linkage (see figure)⁷⁻¹⁰. The amount of Lp(a) in plasma varies from undetectable up to 100 mg dl⁻¹. In extreme cases, 20 per cent of the plasma LDL can be so adorned. Excess LDL is always available, however, so there is no correlation between the concentration of Lp(a) and

LDL in blood.

High levels of Lp(a) are strongly associated with atherosclerosis, as revealed by at least a dozen studies since 1972^{7,11,12}. When the level is above 30 mg dl⁻¹, as it is in about 20 per cent of people, the relative risk of coronary atherosclerosis rises about twofold. When LDL and Lp(a) are both elevated, the relative risk rises to the range of fivefold¹².

The mechanism by which Lp(a) accelerates atherosclerosis is obscure, but a tantalizing clue has emerged from the work reported in this issue³. Through protein sequencing and cDNA cloning, the authors show that apo(a) is a deformed relative of plasminogen, the precursor of the proteolytic enzyme plasmin, which dissolves fibrin clots. Plasminogen, a protein of 791 amino acids, contains 5 cysteine-rich sequences of 80–114 amino acids each, called kringles, followed by a serine protease domain. Each kringle contains three internal disulphide bridges, producing a pretzel-like structure that resembles a Danish cake called *kringle*.



reverse-transcribed as though it contained a uracil instead of the expected cytosine at position 6,457 in the nucleotide sequence (relative to initiator AUG at +1). This substitution changes codon 2,153 from CAA (glutamine) to the termination codon UAA. Oligonucleotide hybridization to genomic DNA confirmed a cytosine at nucleotide 6,457 in the apo B gene of intestinal cells, just as in liver. Direct sequencing of the intestinal apo B mRNA demonstrates that this cytosine is replaced by a uracil during or after transcription⁷. The intestine must possess a highly specific enzyme that modifies a single nucleotide, cytosine, at a single position, 6,457, in a single mRNA. The simplest mechanism would be deamination at the 6-position, converting cytosine to uracil.

One reason that the intestine has gone to the trouble of evolving a new enzyme system to shorten its apo B seems clear. If apo B lacks its carboxy-terminal segment, it cannot bind to LDL receptors. The catabolism of chylomicrons becomes

Kringles are also found in other proteases of the coagulation system, including tissue plasminogen activator (TPA) and prothrombin¹³. In plasminogen, the kringles promote the binding to the substrate, fibrin¹⁴. Plasminogen is proteolytically inactive until it has been cleaved by TPA, or an equivalent enzyme, which breaks the chain at a single arginine residue to form active plasmin.

Long-sought link

Apo(a) contains a hydrophobic signal sequence for secretion followed by 37 copies of kringle 4 of plasminogen, followed by kringle 5 and the protease domain, all highly conserved with respect to plasminogen³. The 36th copy of kringle 4 is altered to contain an extra unpaired cysteine, the likely site of disulphide linkage with apo B-100. Although the amino acids required for protease activity are intact, apo(a) is not a protease because the arginine at the cleavage site for TPA is changed to serine³.

These dramatic findings may provide the long-sought link between lipoproteins and the clotting system. There is evidence, still controversial, that microthrombi on the vessel wall become incorporated into atherosclerotic plaques. Fibrinogen and fibrin are found in these plaques roughly in proportion to cholesterol, and Lp(a) is also present¹⁵. Kringle 4 binds to fibrinogen, although somewhat weakly¹⁴. A protein with 37 copies of this kringle might bind to fibrinogen or fibrin under certain conditions *in vivo*, even though such binding has not yet been detected *in vitro*³. When Lp(a) insinuates itself into the arterial wall following endothelial damage, it may adhere to fibrin, forming a complex that becomes stuck in the wall. Lp(a) might also block proteolysis of fibrin by inhibiting the activation of plasminogen by TPA. The lipoprotein might even compete with plasminogen for access to fibrin, thereby inhibiting cleavage of fibrin clots. In a recent, tantalizing paper¹⁶, a Swedish consortium reported that patients who had a recurrence of myocardial infarction within 3 years had high plasma levels of a 'fast-acting' inhibitor to TPA when compared with patients who had no recurrence. These patients also had a marked elevation in Lp(a), raising the possibility of a relation, possibly indirect, between Lp(a) and TPA inhibition.

Despite these provocative possibilities, the Genentech group, which has considerable experience with TPA, reports³ an inability to find an effect of Lp(a) on the thrombolytic system. These negative findings must be interpreted with caution because they reflect *in vitro* conditions. It is hard to imagine that nature is only teasing us and that the resemblance of apo(a) to plasminogen has no functional consequence.

Many other questions are brought into focus by these findings. For example, what is the nature of the genetic variation in apo(a) among individuals? In careful family studies, Utermann *et al.*¹⁰ recently reported that apo(a) is encoded by a single locus with multiple alleles, many of which produce proteins of different size. It seems likely that the alternative forms result from homologous recombination that adds or eliminates kringles, raising the possibility of functional as well as quantitative differences in apo(a) among individuals.

Utermann *et al.*¹⁰ also find apparent null alleles. These may reflect the usual types of null mutation, or they may reflect the loss of the unpaired cysteine in the 36th kringle, which would prevent attachment to apo B-100. No one knows whether apo(a) would be secreted into blood if it were not attached to apo B-100, and no one knows how long such unattached apo(a) would remain in the circulation. Sera of patients with abetalipoproteinemia should be examined. The specific linkage of apo(a) to apo B-100 suggests an affinity between these two proteins that sets the stage for the disulphide bond. Is it possible that plasminogen also binds to apo B-100 and that this contributes to the atherogenicity of apo B-100 itself?

Selective advantage?

The evolutionary significance of apo(a) is open to question. Through comparisons with plasminogen, McLean *et al.*³ suggest that the apo(a) gene arose about 40 million years ago, about the time of the divergence of Old World and New World monkeys. Indeed, apo(a) has not been found in species lower than primates⁷, although this search must now be repeated in view of possible earlier confusion with plasminogen.

Does apo(a) impart any selective advantage to higher primates, or is it simply a manifestation of a random DNA duplication that has not had sufficient time to be eliminated? Having a form of LDL that binds to fibrinogen might be of selective advantage in wound healing, as the cholesterol-rich particle would concentrate at sites of tissue injury. Macrophages are known to secrete reducing agents, including cysteine and glutathione, in response to a phagocytic challenge¹⁷. This might reduce the disulphide bond between apo(a) and apo B-100, thus liberating the LDL to bind the LDL receptors of proliferating fibroblasts, thereby providing cholesterol.

A final issue relates to the apparent additivity between Lp(a) and LDL in producing atherosclerosis. The frequency of high apo(a) levels is about the same among Japanese people as it is in Caucasians¹⁸. Yet atherosclerosis in Japan was much less frequent than in the United States until the Japanese began to

consume a high-cholesterol diet and developed elevated LDL levels. Armstrong *et al.*¹² have demonstrated an additive effect between LDL and Lp(a) in producing angiographically detectable coronary artery disease.

It seems, therefore, that Lp(a), like the other risk factors for atherosclerosis, is not toxic unless the LDL level is above a certain threshold. Drugs that lower the LDL level may alleviate the toxic effects of Lp(a) without affecting the Lp(a) level itself. Bile-acid binding resins, which lower LDL levels by increasing LDL receptors, do not affect Lp(a) levels¹⁹, presumably because Lp(a) is a relatively poor ligand for the LDL receptor²⁰. Other drugs, such as neomycin and nicotinic acid, which lower LDL by other mechanisms, do lower Lp(a)²¹. The effects on Lp(a) of the newest LDL-lowering drugs, HMG CoA reductase inhibitors, must now be examined.

The three papers we have discussed here represent technically demanding feats of molecular biology. In particular, the work with apo(a) demanded that the Genentech team pick its way through a minefield of potential artefacts emanating from a cDNA with 37 copies of a highly repetitive sequence, all very similar to sequences in another family of genes. Clinicians draw on the powers of molecular biology to unmask human disease. Occasionally, they repay their debt to the molecular biologists by providing a glimpse into basic biological processes — as is the case with the work discussed here. □

1. Powell, L.M. *et al.* *Cell* **50**, 831–840 (1987).
2. Chen, S.-H. *et al.* *Science* **238**, 363–366 (1987).
3. McLean, J.W. *et al.* *Nature* **330**, 132–137 (1987).
4. Kane, J.P. *A. Rev. Physiol.* **45**, 637–650 (1983).
5. Milne, R.W. & Marcel, Y.L. *Can. J. Biochem. Cell Biol.* **63**, 906–912 (1985).
6. Ludwig, E.H. *et al.* *DNA* **6**, 363–372 (1987).
7. Kostner, G.M. in *Low Density Lipoproteins* (eds Day, C.E. & Levy, R.S.) 229–269 (Plenum, New York, 1976).
8. Fless, G.M., ZumMallen, M.E. & Scanu, A.M. *J. biol. Chem.* **261**, 8712–8718 (1986).
9. Gaubatz, J.W., Chari, M.V., Nava, M.L., Guyton, J.R. & Morrisett, J.D. *J. Lipid Res.* **28**, 69–79 (1987).
10. Utermann, G. *et al.* *J. clin. Invest.* **80**, 458–465 (1987).
11. Dahlen, G., Ericson, C., Furberg, C., Lundkvist, L. & Svärdsudd, K. *Acta med. Scand.* **531** (suppl. 6), 11–15 (1972).
12. Armstrong, V.W. *et al.* *Atherosclerosis* **62**, 249–257 (1986).
13. Patthy, L., Terxler, M., Vali, Z., Banyai, L. & Varadi, A. *FEBS Lett.* **171**, 131–136 (1984).
14. Trexler, M., Vali, Z. & Patthy, L. *J. biol. Chem.* **257**, 7401–7406 (1982).
15. Walton, K.W., Hitchens, J., Magnani, H.N. & Khan, M. *Atherosclerosis* **20**, 323–346 (1974).
16. Hamsten, A. *et al.* *Lancet* **ii**, 3–8 (1987).
17. Rouzer, C.A., Scott, W.A., Griffith, O.W., Hamill, A.L. & Cohn, Z.A. *J. biol. Chem.* **257**, 2004–2008 (1982).
18. Murai, A., Miyahara, T., Fujimoto, N., Matsuda, M. & Kameyama, M. *Atherosclerosis* **59**, 199–204 (1986).
19. Vessby, B., Kostner, G., Lithell, H. & Thomas, J. *Atherosclerosis* **44**, 61–71 (1982).
20. Armstrong, V.W., Walli, A.K. & Seidel, D. *J. Lipid Res.* **26**, 1314–1323 (1985).
21. Gurakar, A., Hoeg, J.M., Kostner, G., Papadopoulos, N.M. & Brewer, H.B. Jr *Atherosclerosis* **57**, 293–301 (1985).

Michael S. Brown and Joseph L. Goldstein are in the Department of Molecular Genetics, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235, USA.

Are high- T_c superconductors doped Mott insulators?

SIR—In a recent report in *Nature*, Sir Nevill Mott¹ pointed out that there are almost as many theories of high- T_c superconductivity, where T_c is the critical temperature, as there are theorists. However, he showed a preference for a phonon-mechanism theory. At the European Physical Society in April 1987, in Pisa, I suggested that high- T_c superconductors are doped Mott insulators, that is, there exist strong Coulomb correlations in the narrow band made by the hole of Cu^{2+} ions. This should lead not only to the insulating behaviour but to an antiferromagnetic ground state for the undoped La_2CuO_4 compound. My theoretical considerations received scepticism except from K. A. Müller.

Since the recent experiments at AT&T Bell Laboratories² and by the Berkeley group which showed no isotope effect in the superconductor $\text{YBa}_2\text{Cu}_3\text{O}_7$, the importance of correlations as a possible mechanism for superconductivity is now recognized. Subsequently, in *Nature*, Anderson and Abraham³ excluded all theories based on a phonon mechanism and argued that experiments indicate a strong short-range interelectronic repulsion U . But the quoted theories are all in the small U limit except for Anderson's resonant valence bond (RVB) model⁴ which starts from a non-magnetic insulator.

My proposal that high- T_c superconductors are doped antiferromagnetic Mott insulators is now supported by several reported magnetic properties of these compounds. But I believe that the importance of these results has been insufficiently stressed. Furthermore, starting from an antiferromagnetic Mott insulator, I present a physical picture which is different from the one developed by Anderson in the RVB model.

As far as the magnetic properties of these materials are concerned, neutron diffraction studies⁵ on La_2CuO_4 have shown antiferromagnetism. Susceptibility measurements⁶ performed at IBM on doped superconducting materials show a large negative Curie paramagnetic temperature, indicating antiferromagnetic correlations in these compounds. For the other family, the compound $\text{YBa}_2\text{Cu}_3\text{O}_{6.5}$, which contains only Cu^{2+} ions, shows semiconducting behaviour and also a susceptibility with a Curie-Weiss component. More experiments are needed but we expect that neutron diffraction above T_c would show directly antiferromagnetic correlations in the two families.

On the theoretical side, I recently used an approach developed for explaining the magnetic properties of a lattice of Jahn-Teller ions⁷. I applied it to Cu^{2+} in a perovskite structure and predicted an

antiferromagnetic ground state for La_2CuO_4 that neutron experiments confirmed. This shows therefore that the usual approach of the lattice of Jahn-Teller ions developed in the large- U limit can be successfully applied to these new copper perovskites. In the large- U limit, the undoped materials can be either an application of Anderson's RVB model or an antiferromagnetic Mott insulator. Present experiments strongly support the second point of view.

A main difference between my approach and that of Anderson's is as follows. While doping (creating holes in the lower Hubbard sub-band) I do not allow the pre-existing pairs of the RVB model to move as they are in an antiferromagnetic state. In fact the antiferromagnetic state is destroyed but the strong interactions between spins on the copper atom will persist. At high temperatures, it leads to the Curie-Weiss behaviour with a negative Curie paramagnetic temperature; at temperatures linked with the strength of the interactions, it leads to superconductivity. The behaviour of T_c as a function of doping is different from the RVB model and so are some of the other physical properties. For example, I calculate a jump in the specific heat function of doping⁸. I predict no pairs above T_c but the properties are described through a correlated hole hopping within a lower Hubbard sub-band as discussed at length by Mott⁹. Below T_c , the antiferromagnetic interactions between carriers are directly responsible for the new type of superconductivity, which is singlet and anisotropic.

M. CYROT

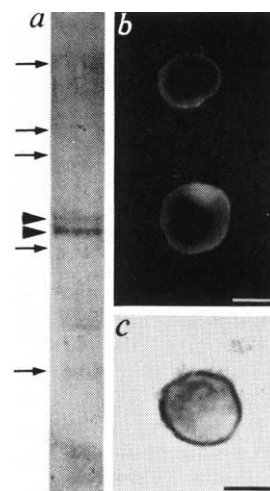
Laboratoire Louis Néel,
CNRS, 25 Avenue des Martyrs,
Grenoble, 166X, 38043 Grenoble,
Cedex, France

1. Mott, N.F. *Nature* **327**, 185 (1987).
2. Batlogg, B. *et al. Phys. Rev. Lett.* **58**, 2333 (1987).
3. Anderson, P.W. & Abrahams, E. *Nature* **327**, 363 (1987).
4. Anderson, P.W. *Science* **235**, 1196 (1987).
5. Yamaguchi, Y. *et al. Jap. J. appl. Phys.* **26**, L447 (1987).
6. Green, R.L. *et al. Solid St. Commun.* (submitted).
7. Cyrot, M. *Solid St. Commun.* **62**, 821 (1987).
8. Cyrot, M. *Solid St. Commun.* (submitted).
9. Mott, N.F. *Metal Insulator Transitions* (Taylor and Francis, London, 1974).

Synapsin or protein 4.1 in chromaffin cells

SIR—The role of the cytoskeleton in the control of exocytosis in secretion from adrenal chromaffin cells and neurotransmitter release from nerve terminals was discussed in our recent News and Views articles^{1,2}. Commenting on these articles, Siegel³ speculates about common mechanisms in the control of exocytosis in chromaffin cells and nerve terminals and poses the question of whether synapsin I is present in chromaffin cells.

Synapsin I is believed to be important in the nerve terminal, where its state of



Presence of a 4.1-like protein in bovine adrenal chromaffin cells. *a*, Immunoblotting with anti-red-cell 4.1 of an homogenate of chromaffin cells purified by differential plating after dissociation of adrenal medulla. Arrowheads, immunoreactive doublet; arrows, positions of relative molecular mass standards of 205,000 (205K), 116K, 94K, 67K and 43K (from top to bottom). Immunocytochemical localization of the 4.1-like protein using affinity-purified anti-red cell 4.1 by *b*, immunofluorescence or *c*, immunoperoxidase staining in chromaffin cells maintained in culture for 24 h showing intense staining at the cell periphery. Scale bars in *b* and *c*, 10 μm . The preparation and characterization of anti-red-cell 4.1 was described previously⁴.

phosphorylation could determine the extent of actin filament crosslinking and possibly linkage of actin filaments to synaptic vesicles⁴. As Siegel³ points out, secretory granules from chromaffin cells do contain a phosphoprotein⁵ of the same relative molecular mass as synapsin I. But, unlike synapsin I, this phosphoprotein does not show clear calmodulin-dependent phosphorylation⁵ and is in any case very much less abundant in chromaffin granules than is synapsin I in synaptic vesicles. Furthermore, denervation of the adrenal medulla results in almost complete disappearance of synapsin I from the gland⁶, suggesting that it is present in nerve terminals in the adrenal medulla but not chromaffin cells, and immunocytochemical studies⁷ fail to detect synapsin I in chromaffin cells.

Synapsin I is functionally similar to, and shows immunological cross-reactivity with, red-cell protein 4.1 (ref. 8), which is involved in linking spectrin and actin filaments to the plasma membrane. Because of the similarity between the suggested role of the cytoskeleton in nerve terminals and chromaffin cells in restricting vesicle access to exocytotic sites before stimulation, we have examined the possibility that a 4.1-like protein is present in chromaffin cells instead of synapsin I. As shown in the figure, there is indeed an immunoreactive doublet of similar molecular mass (about 70,000) to red-cell

4.1 in chromaffin cells, and immunofluorescence and immunoperoxidase staining with affinity-purified anti-red cell 4.1 shows that the chromaffin cell-related protein is concentrated in the cell periphery, associated either with the plasma membrane or with the cortical cytoskeleton. Therefore, the chromaffin cell form of protein 4.1 could be a target for protein kinases in the regulation of the peripheral cytoskeleton of the chromaffin cell, in particular of the association of cortical actin filaments with fodrin and the plasma membrane.

We agree with Siegel's suggestion that there are likely to be common molecular mechanisms in the control of exocytosis in chromaffin cells, nerve terminals and also other cells. In the case of the regulation of the interaction of actin filaments with fodrin and membranes it seems that similar control mechanisms may occur, but with different members of the same functional family of molecules as targets in different cell types.

ROBERT D. BURGOYNE
MRC Secretary Control Group,
The Physiological Laboratory,
University of Liverpool,
PO Box 147,
Liverpool L69 3BX, UK

ANTHONY J. BAINES
Biological Laboratory,
University of Kent,
Canterbury,
Kent CT2 8NJ, UK

1. Burgoyne, R.D. & Cheek, T.R. *Nature* **326**, 448 (1987).
2. Baines, A.J. *Nature* **326**, 646 (1987).
3. Siegel, D. *Nature* **327**, 467 (1987).
4. Bahler, M. & Greengard, P. *Nature* **326**, 704–707 (1987).
5. Burgoyne R.D. & Geisow, M.J. *FEBS Lett.* **131**, 127–131 (1981).
6. Fried G. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2717–2721 (1982).
7. Navone, F. *et al. J. Cell Biol.* **103**, 2511–2527 (1986).
8. Baines, A.J. & Bennett, V. *Nature* **315**, 410–413 (1985).

Artificial guide stars for adaptive imaging

SIR—The experiments of Thompson and Gardner¹ followed proposals by Foy and Labeyrie² to create artificial 'guide stars' for adaptive imaging in astronomy by the backscattering from the upper atmosphere of laser pulses projected from a telescope. It has been suggested that if sufficient energy could be returned from a thin scattering layer it could provide the information required to drive adaptive optical systems able to compensate for the atmospheric degradation of images of faint objects of astronomical interest.

There appear to be difficulties with this proposal, as in perhaps the simplest possible case the backscattered light provides no information at all about refraction in the lower atmosphere.

Suppose, for example, that the scattering layer is very distant and that the aperture of the telescope/projector may be completely covered by a thin transparent

wedge whose only effect is to refract light passing through it through a small angle. The image of a distant real object will move if this wedge is introduced, but an image formed by backscattering will be unaffected because the incoming beam will precisely retrace its outgoing path.

Complexities introduced by random variations of refraction across single or multiple apertures do not seem likely to improve this situation, and it therefore seems improbable that techniques of this kind will yield the benefits that have been suggested.

J.D.H. PILKINGTON
Royal Greenwich Observatory,
Hailsham,
East Sussex BN27 1RP, UK

1. Thompson, L.A. & Gardner, C.S. *Nature* **328**, 229–231 (1987).
2. Foy, R. & Labeyrie, A. *Astr. Astrophys.* **152**, L29–L31 (1985).

THOMPSON AND GARDNER REPLY—There are three potential problems that might cause a laser guide star to move relative to natural background stars: laser jitter, non-sidereal motion of the telescope, and low frequency atmospheric disturbances.

Problems of laser jitter can be removed by monitoring the laser beam as it leaves the telescope, and any detected motion can be nulled in the computer which analyses the wave front error. Telescope jitter can be removed by locking the instrument onto bright natural stars; these stars need not be in the isoplanatic patch.

The problem of low frequency atmospheric disturbances is more subtle, but it too can be corrected. If the natural guide stars used to compensate for telescope jitter are spaced over a large area in the sky (5° to 10° patch), the average low frequency tilt will be zero. Then if we simultaneously monitor the motion of a natural star near the laser target, its motion will show evidence of all low frequency disturbances occurring at the target. This signal can also be sent to the computer which analyses the wave front error, and the computer can make the required corrections.

Even if the multiple guide star technique fails to remove all low frequency wave front errors, there is a less acceptable but safe fall-back solution. A moderately fast data recording system can be used at the final focal plane, and the adaptively sharpened images can be recorded with relatively short exposure times. These short exposures can be recentred and coadded *ex post facto*, thereby giving a sharpened and stable image.

LAIRD THOMPSON
CHET GARDNER
Department of Astronomy,
University of Illinois at
Urbana-Champaign,
1011 West Springfield Avenue,
Urbana, Illinois 61801, USA

Red Queen versus Tangled Bank models

SIR—Burt and Bell¹ present an analysis of chiasma frequency in mammals, which they claim supports the theory that genetic recombination is maintained by selection pressures resulting from host–pathogen interactions, which cause cyclical fluctuations in the nature of selectively favoured gene combinations (the Red Queen hypothesis²). The basis for this claim is that the excess of chiasma frequency per individual over haploid chromosome number is positively correlated with age to maturity (which they equate with generation time); they assert that this correlation is predicted by the Red Queen model, but not by the only other model that they consider to be a viable alternative—the model of sib-competition in a spatially varying environment^{3,4}, which they term the Tangled Bank model.

Their reasoning is that the timescale for interactions between mammals and their pathogens is such that species with a long generation time will be more likely to encounter a change in selection pressure over a small number of generations, compared to species with short generations, and hence will be more likely to evolve high rates of recombination. They also claim that species with large litter size will be more likely to experience sib-competition; because these tend to have short generations, the Tangled Bank model predicts a negative correlation with generation time.

There are several difficulties with this study. In the first place, the authors' treat each species as an independent data point in their correlation tests. It is a well-known principle of the comparative method that such correlations are biased in the direction of spuriously positive results if there is a lack of evolutionary independence between groups of data points⁵. Inspection of Burt and Bell's Fig. 1 shows that there are many points contributed by species of primates and rodents; much of their correlation is evidently due to the fact that rodents tend to have low chiasma frequencies per bivalent and primates high ones. It is not at all clear what this means, in view of the many biological differences between these groups. It could, for example, be the case that sib-competition is more intense in the highly social primates, leading to a greater advantage to high recombination through the Tangled Bank model. Such sib-competition has been invoked as a cause of female-biased sex ratios in primates⁶.

Second, their rejection of all possible explanations except the Red Queen and Tangled Bank models, on the basis of Bell's survey² of the comparative data on the taxonomic distribution of asexuality in animals, is not accepted by many workers in the field. As pointed out in at least

one review of Bell's book, many of the correlations reported by Bell could be explained in terms of the reproductive advantage enjoyed by asexual variants in low-density habitats⁷. Thus, this type of study may not shed much light on the nature of the forces that actively maintain recombination. Furthermore, a new theory for the advantage of sex and recombination, based on the maintenance of deleterious genes by mutation pressure, has since been proposed by Kondrashov^{8,9}, and was not considered in Bell's original survey⁷.

If the somewhat arbitrary procedure of considering only the Red Queen and Tangled Bank models is discarded, then the interpretation of any real correlation between generation time and rate of recombination becomes hard. Theories of recombination that are in part dependent on the effects of finite population size, such as Muller's ratchet and the Fisher-Muller model of selection for favourable genes³, would often predict a higher advantage to increased recombination in populations with smaller effective sizes. Large mammals tend to have smaller numbers of individuals in the species, as well as larger generation times, consistent with the claimed correlation. Similarly, the Fisher-Muller model, the related model of Strobeck *et al.*³, and my own model of temporally fluctuating environments¹⁰, all predict a positive correlation between generation time and rate of recombination.

In summary, the claim by Burt and Bell to have found strong support for the Red Queen model of the evolutionary advantage of recombination is without firm theoretical or observational basis.

BRIAN CHARLESWORTH

Department of Biology,
The University of Chicago,
1103 East 57th Street,
Chicago, Illinois 60637, USA

1. Burt, A. & Bell, G. *Nature* **326**, 803–805 (1987).
2. Bell, G. *The Masterpiece of Nature: the Evolution and Genetics of Sexuality* (Croom Helm, London, 1982).
3. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
4. Bell, G. *Experientia* **41**, 1235–1245 (1985).
5. Felsenstein, J. *Am. Nat.* **125**, 1–15 (1985).
6. Clark, A.B. *Science* **201**, 163–165 (1978).
7. Wilson, D.S. & Gleeson, S.K. *Evolution* **37**, 428–430 (1983).
8. Kondrashov, A.S. *Genet. Res.* **40**, 325–332 (1982).
9. Kondrashov, A.S. *Genet. Res.* **44**, 199–217 (1984).
10. Charlesworth, B. *Genetics* **83**, 181–195 (1976).

SIR—Burt and Bell¹ showed a correlation between chiasma frequency and age at maturity in mammals. They interpret this as support for the Red Queen hypothesis, that high recombination is advantageous when there are large changes in the environment between generations, which change the phenotypic optimum. They assume that the extent of environmental change between generations is greatest in those species that have the longest generation intervals. I believe that this assumption is weak and that the

correlation could well have another explanation.

If parasites form the relevant part of the environment, as suggested^{1–3}, then there may be a correlation between environmental change and generation length. However, because parasites with long generations tend to inhabit hosts with long generations, the mean generation length of parasites will be rather larger in those hosts that have long generations, so the rate of parasite evolution per unit of time will be lower than in parasites of hosts with short generations. If organisms other than parasites are important, the correlation will be weaker because animals tend to interact more with species of a similar size than with species of a dissimilar size and size is correlated with generation length. If the physical environment is the important component, the correlation between age at maturity and environmental change may be the opposite to that assumed. Physical factors change more markedly between seasons than from year to year, so species that have several generations per year suffer greater degrees of environmental change between generations than do species that have an annual cycle. Species with generation lengths of several years may tend to suffer higher levels of environmental change than those with annual cycles because of long-term changes, but the effects of such long-term changes will generally be overwhelmed by the law of large numbers: the mean environment in two successive years is likely to differ more than the mean in two successive spans of ten years.

The Tangled Bank hypothesis is that high recombination is advantageous when there is strong selective competition between sibs for survival, with some environmental change between generations. Burt and Bell state that if this hypothesis is true, one would expect to find a correlation between chiasma frequency and fecundity, which they did not find. But this rests on the view that sib-competition is stronger in species with high fecundity, which may be false. If animals produce litters that maximize the expected number of surviving offspring, litter size will reflect the resources available and inter-specific differences in the degree of sib competition will be unrelated to inter-specific differences in litter size.

Furthermore, so long as there is some environmental change between generations, recombination is favoured by any intense competition, not only sib-competition. Thus, an alternative explanation for the observed correlation is that competition is greater in longer-lived animals, which tend to be K-selected. This explanation, involving both environmental change and competition, reminds us that the Red Queen and the Tangled Bank are extremes of a continuum involving these two processes.

Correlations of this sort are valuable contributions to our understanding of selective forces moulding living organisms. However, the relationships between life history characteristics are complex, so most correlations will be explicable by more than one hypothesis. In the present case, the search for a general conclusion may never be successful: the way in which chiasma frequency is moulded may be different in different species; changes in the environment between generations may be more important for some, the strength of competition between individuals for others. To focus attention on one or the other may help to clarify ideas but must not constrain us into thinking that only one is important.

JEREMY J.D. GREENWOOD

Department of Biological Sciences,
Dundee University,
Dundee DD1 4HN, UK

1. Burt, A. & Bell, G. *Nature* **326**, 803–805 (1987).
2. Jaenike, J. *Evol. Theory* **3**, 191–194 (1978).
3. Hamilton, W.D. *Oikos* **35**, 282–290 (1980).

SIR—In their recent letter (*Nature* **326**, 803–805, 1987) Burt and Bell may be too optimistic. The correlations they find between mammalian chiasma frequencies and various life history parameters do not necessarily support the Red Queen or reject the Tangled Bank theories for at least three reasons. First, the Tangled Bank has not yet been formulated as a theory of recombination, but as a theory concerning the maintenance of sexual reproduction. It thus does not predict any particular level of recombination, but only whether any level of recombination is favoured over clonal reproduction, that is, over no recombination. Hence the correlations with recombination rate to be expected for the Tangled Bank hypothesis are not clear. Second, even if it is assumed that the Tangled Bank is a theory of recombination, the correlations expected by Burt and Bell are not obvious. It could be argued that the organisms with short generation times and high litter sizes are not under severe competition. The Tangled Bank would then predict low recombination. On the other hand, organisms with long generation times may be selected by competitive effects, and are therefore predicted to have a higher rate of recombination. Hence the Tangled Bank would predict the positive correlation between recombination rate and generation time shown by Burt and Bell. Furthermore, it is well known that in host-parasite systems it is not the absolute generation time that affects the outcome of selection, but the differences in generation times of host and parasite. Thus if generation times of host and parasite themselves are correlated, the predictions of the Red Queen could range from positive to strongly negative correlations of recombination rate with generation time. Third, the

results of Burt and Bell would be strengthened if a partial correlation of recombination rate with age at maturity and litter size were performed, and if the effects of common ancestry were considered. It seems that species within a family have fairly constant recombination rates, despite large differences in generation times. However, this latter point might not be very relevant, because the data on domestic animals and other, previously published, data show that recombination rates can change rapidly.

JACOB C. KOELLA

Zoologisches Institut der Universität,
Rheinsprung 9, CH-4051 Basel,
Switzerland

BURT AND BELL REPLY—We have argued¹ that the positive correlation between excess chiasma frequency (EC) and age to maturity among mammals supports the Red Queen theory of recombination, according to which longer-lived hosts must have higher rates of genetic recombination per generation to escape from the counter-adaptations of their parasites. It was an implicit assumption of our test that the number of parasite generations per host generation, and therefore the relative rate of parasite evolution, is positively correlated with host generation time. Greenwood and Koella question this assumption, suggesting that the generation time of parasites may increase isometrically with that of their hosts. Unfortunately, neither author provides any data: we have therefore collated measurements of the prepatent period (time from infection of definitive host to release of first propagules into the external environment) of 37 parasitic species in four major taxa and the age to maturity of their mammalian hosts.

There is indeed a very highly significant positive association between parasite

prepatent period and host age to maturity ($P < 0.001$), with no indication of heterogeneity of slopes among the four parasite taxa. However, the common slope of this association is much less than unity and so the number of parasite generations per host generation increases steeply with host generation time (Fig. 1). This result is completely consistent with our earlier interpretation and decisively refutes Greenwood's and Koella's conjectures.

We have further argued that the negative correlation between EC and litter size among mammals falsifies the Tangled Bank theory. Greenwood and Koella dispute our rejection, suggesting that longer-lived and less fecund species are subject to more intense competition. This claim, again made without empirical support, misses the basic proposition of the Tangled Bank: that the increased phenotypic variance created by recombination functions to reduce the level of competition within families. It is emphatically a theory of competition between siblings; claims that it does not predict a positive correlation with litter size contradict both common sense and conclusions from computer simulations². Indeed, such a theory clearly does not predict that mammals with litter sizes of one will be among those with the highest levels of recombination, yet this is precisely what we observe. Furthermore, it would appear that litter size is not even a minor contributor to the observed variance in EC: the partial correlation, after removing the effects of age to maturity, is insignificant ($r = -0.320$, d.f. = 21, $0.2 > P > 0.1$). This result seriously damages Greenwood's concluding plea for plurality.

Charlesworth is not interested in saving the Tangled Bank, but rather several theories not considered in our original paper. However, it is not clear why — he offers nothing to counter the well-

documented shortcomings of all but one of these theories^{3,4}. The author of the remaining theory⁵ concludes from simulations that his model cannot account for crossing over in organisms with many chromosomes; 70% of our mammals have haploid numbers greater than any of those used in his simulations ($n > 16$).

Finally, Charlesworth claims that our treatment of each species' EC as statistically independent is likely to give spurious positive results. His argument conflates two separate issues. The first, common ancestry and phylogenetic inertia, is unimportant when there exists additive genetic variance for the character under investigation, for the response to selection will be instantaneous on a palaeontological timescale. Evidence for such variance in EC comes both from artificial selection and heritability studies^{3,4} and from the twofold increase in EC among domesticated mammals¹. Thus the EC of each species is an independent product of natural selection and should be analysed accordingly. Nevertheless, to allay any possible doubts, we have reanalysed our data at the family level; the correlation of family means is almost identical with that of species and the null hypothesis is rejected at the same level that we reported ($r = +0.883$, d.f. = 10, $P < 0.001$).

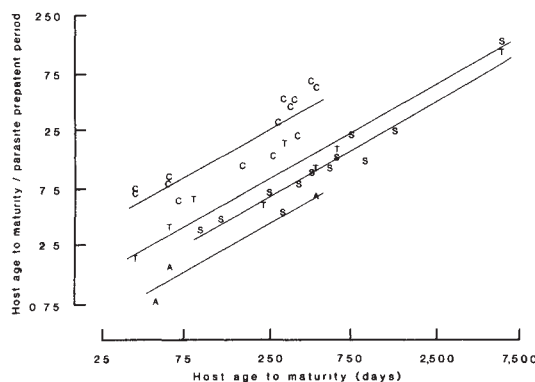
The second issue is that of confounding selection pressures, for although EC is correlated with age to maturity, it is presumably also correlated with many other characters related to age to maturity. This problem motivated our partial correlation analysis of EC and age to maturity, keeping body weight constant. The latter is strongly correlated with age to maturity among mammals ($r = +0.844$ in our data set), but when the three variables are analysed simultaneously, the partial correlation of EC and age to maturity remains large and significant ($r = +0.689$, $P < 0.001$), whereas that of EC and body weight becomes insignificant ($r = +0.021$, $P > 0.9$).

AUSTIN BURT*
GRAHAM BELL

Department of Biology,
McGill University,
1205 Avenue Docteur Penfield,
Montreal, Quebec H3A 1B1,
Canada

*Present address: Department of Zoology, South Parks Road, Oxford OX1 3PS, UK.

Fig. 1 The timescale of mammalian parasites. The ratio of host age to maturity to parasite prepatent period increases with host age to maturity in four parasite taxa with mammalian definitive hosts: A, Acanthocephala⁶; S, Schistosomatiidae⁷; T, Cestoda⁸; and C, Coccidia⁹ (note logarithmic axes). When a parasite infests more than one species of host, or vice versa, means were used. To quantify the within-taxon association, the data set was transformed such that the four groups



had equal bivariate means. The transformed data set has Pearson correlation coefficient $r = 0.948$ ($n = 37$) and reduced major axis slope $v_{yx} = 0.867$. This latter value is the common slope of the within-taxon association, $\bar{v}$, and has standard error about equal to that calculated by analysis of covariance: $s_{\bar{v}} = 0.049$. Both major axis regression and analysis of covariance showed the four slopes to be homogeneous ($0.9 > P > 0.5$). The four solid lines are drawn with slope $\bar{v}$ through the (untransformed) bivariate mean of each parasite taxon.

1. Burt, A. & Bell, G. *Nature* **326**, 803–805 (1987).
2. Maynard Smith, J. *J. theor. Biol.* **63**, 245–258 (1976).
3. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
4. Bell, G. *The Masterpiece of Nature: the Evolution and Genetics of Sexuality* (Croom Helm, London, 1982).
5. Kondrashov, A.S. *Genet. Res.* **44**, 199–217 (1984).
6. Crompton, D.W.T. in *Biology of the Acanthocephala* (eds Crompton, D.W.T. & Nickol, B.B.) 213–271 (Cambridge University Press, 1985).
7. Loker, E.S. *Parasitology* **87**, 343–369 (1983).
8. Kennedy, C.R. in *Biology of the Eucestoda* Vol. 1 (eds Arme, C. & Pappas, P.W.) 27–80 (Academic, London, 1983).
9. Khayson, Y.M. *Life Cycles of Coccidia of Domestic Animals* (University Park, Baltimore, 1972).

Different interpretations

Bernard Wood

Fossils, Teeth and Sex: Perspectives on Human Evolution. By Charles E. Oxnard. Hong Kong University Press/University of Washington Press: 1987. Pp.281. US\$38 (Far East), US\$35 (North America), £28 (UK).

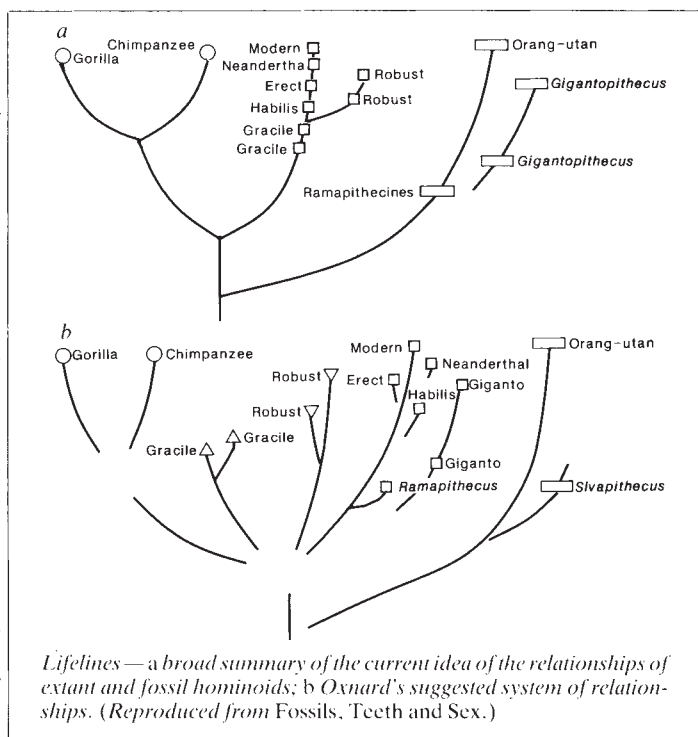
THIS new book of Charles Oxnard's is a mixture of monograph and compendium. It is the fourth of a series, but in size, style and content, it comes closer to the third volume, *The Order of Man*, than to the first two. In it, the author tackles the topic of sexual dimorphism in primates, and in particular its evolutionary and behavioural implications.

Sexual dimorphism is of especial interest to primate biologists and palaeontologists for two reasons. First, if you know what to look for in a fossil, it is possible to detect evidence of consistent skeletal and dental differences between the sexes. Secondly, because of its apparently strong, if complicated, associations with social and locomotor behaviour, the expression of sexual dimorphism in a taxon has a predictive value for those who are eager to put flesh on the bones of the fossil record. Although sexual dimorphism is the main subject of the book, it does not exclusively concern itself with that topic. For, in the course of the text, Oxnard takes the opportunity to comment on the work of practitioners of cladistics, classical morphology and multivariate morphometrics.

Oxnard's thesis is essentially a simple one. He claims, with justification, to have demonstrated in a series of papers that there are several types of sexual dimorphism. He has shown, and I believe it is a novel demonstration, that there are differences in variance between the sexes of some primate taxa, notably the extant African apes. Thus, when looking at the sex distributions of a variable, we should not only take note of the mean values, but also the shape and area of the bell-curves. Oxnard claims that the area of each curve provides a clue to the sex ratio of the population which the fossil collection samples. The author sensibly focused his study of sexual dimorphism on dental evidence; this was a pragmatic decision prompted by several factors, not the least of which is that teeth are preferentially preserved in fossil assemblages.

The patterns shown by the extant higher primate taxa (*Homo*, *Pan*, *Gorilla* and

Pongo) provide the reference set for the analysis of three sets of fossil data. The first is from a large collection of primate teeth of Late Miocene age from Lufeng, China. The second is a similar-sized sample (more than a thousand) of *Gigantopithecus* teeth from three cave sites in China. The third set comprises measurements of tooth crowns from fossil hominids ranging from the Pliocene *Australopithecus afar-*



Lifelines—a broad summary of the current idea of the relationships of extant and fossil hominoids; b Oxnard's suggested system of relationships. (Reproduced from *Fossils, Teeth and Sex*.)

ensis to the much later Neanderthals. The Lufeng teeth are treated as two, taxonomically distinct, samples, designated *Ramapithecus* and *Sivapithecus*, and the apparent differences in the pattern of sexual dimorphism between these two samples is the basis for much of the subsequent analysis and interpretation.

In the final section of the book Oxnard presents the evidence for, and the consequences of, a radical interpretation of hominid evolution. He proposes that *Ramapithecus* and all *Homo* taxa are linked together (see figure) in a lineage characterized by a 1:1 sex ratio and equal male and female dispersions combined with only small to moderate size differences between the sexes. Because their variances and sex ratios are unequal, and the size differences between the sexes large, all australopithecines are thus

excluded from this lineage, so that any features which have hitherto been cited as characteristic of hominids as a whole must now be regarded as convergent or primitive characters. Oxnard concludes the book with a section in which he discusses the behavioural, adaptive and even sociological implications of these proposals.

In many ways *Fossils, Teeth and Sex* is a *tour de force*. Its scope is broad and the standard of written exposition is uniformly high. It is, however, the task of the reviewer to offer readers the equivalent of a structural survey of a house. As impressive as a building is as an edifice, the reviewer has to probe the foundations and the design to see if they are sound.

The extant comparative evidence is clearly set out in the first part of the book. Oxnard deserves recognition for drawing

our attention to the potential richness of information subsumed under the deceptively simple heading of 'sex differences'. I would however take issue with his interpretations of the field 'before Oxnard'. Despite the interpretations, and indeed quotations, ascribed to us, several workers had previously drawn attention to the considerable shape differences between some male and female primates and had provided quantitative evidence of the higher proportion of shape difference which occurs between the skeletons (and dentitions) of forms, such as male and female chimpanzees, which differ little in overall body size. Oxnard acknowledges the taphonomic biases which may vitiate his conversion of fossil assemblages data into population sex ratios, but his approach is innovative and laudable. Nonetheless the key-

stone of his logical edifice is his interpretation of the Lufeng material, and here there are at least two problems.

The first is one of simple logic. Why, for example, is the author prepared to accept sexual dimorphism as the explanation for the variation within the *Gigantopithecus* sample, even though he describes it as a "remarkable situation" (p.153), "totally different from any other form, even the most dimorphic of the large apes, *Gorilla*" and "totally different from all other living and fossil species" (p.166), and yet dismiss the same explanation for the no less extreme sex differences in the Lufeng sample as "utterly unlikely" (p.93) and "completely untenable" (p.113)?

It seems clear that Oxnard is prepared to make this exception because Professor Wu Rukang and his colleagues had made prior decisions about the presence of two

taxa at Lufeng, and the author rightly credits his Chinese colleagues as being "far more experienced with this particular group of ramapithecines than workers anywhere else in the world" (p.92). It is thus unfortunate that these same scientists in a paper published last year, now conclude that sexual dimorphism is the most likely explanation for the observed variation at Lufeng.

Secondly, Oxnard draws comfort from the fact that each of the Lufeng subsets shows univariate distributions and multivariate patterns that could be interpreted as sexual dimorphism. But pattern is not proof, and his Fig. 5.6 is a chastening reminder of that. In that figure it appears as if the 'robust' australopithecines are dimorphic in canine breadth; indeed the shape of the histogram is familiar from the extant studies. However, the two 'sexes' are in fact geographical subsets of australopithecines. This is not to say that Oxnard's interpretation is wrong. But it does suggest that, to return to the metaphor, there is loose mortar in the foundations and there are some flaws in the design.

There are also one or two problems when Oxnard turns his attention to the fossil hominids. Although I am pleased to see a talus bone from Koobi Fora playing such a crucial role in the argument, the one he refers to, KNM-ER 813, is most unlikely to be older than 1.8 million years. Thus, it will have to be denied its chance to star as a 4.5-million-year-old example of

Homo. Elsewhere in the book, I was intrigued by Oxnard's attitudes to cladistics. At the very beginning of the text, he adopts a 'cladistic' stance and, quite properly, warns against the unthinking pursuit of ancestors. Yet later on cladistics is roundly criticized, and yet later again his summary diagram will please cladists because it is 'ancestor-free'.

The final comments on the book must concern the 'new picture of human evolution' that is its culmination. It is indeed a radical and brave proposition to suggest that australopithecines and hominines (that is, *Homo*) have no exclusive common ancestor. It is not such a bizarre notion as it seems; but the proposer does need to explain how the two groups come to share the characters which we presently recognize as hominid, for example reduction of the canine teeth and relative brain-size enlargement. There also have to be explanations for the characters which are apparently shared between *Australopithecus africanus* and *Homo*. However outrageous a scheme it may seem to some, it is, nonetheless, an alternative which must be considered. Cladistic studies of hominids all suggest that convergence was more common than many palaeontologists had thought. Whether, however, it is as rife as Oxnard implies in his new system remains to be seen. □

Bernard Wood is Derby Professor in the Department of Human Anatomy and Cell Biology, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK.

Simulating life's work

Michael Levitt

Dynamics of Proteins and Nucleic Acids.

By J. Andrew McCammon and Stephen C. Harvey. Cambridge University Press: 1987. Pp. 234. £27.50, \$39.50.

NINETEENTH-century philosophers would have loved this book. Starting from atoms, which are assumed to interact through simple forces and move by classical mechanics, those using molecular dynamics aim to simulate nature in all her glory, calculating with reliability the measurable physical and chemical properties of important biological molecules such as proteins and nucleic acids. Computer simulation of biological macromolecules encompasses the complete chain of reductionist thought, working up the pyramid of science from arithmetic and mechanics, through physics and chemistry, to the prediction and understanding of biological phenomena. In their preface McCammon and Harvey express their pleasure in "using Newton's laws to

interpret basic events in biochemistry".

What is truly amazing is that such an approach, which is almost impertinent in its simplification and scope, actually works. What is more, simulations lasting very short periods of time on systems that are ridiculously small by normal standards give reasonable results. Thus, 200 water molecules simulated for 10^{-11} seconds give to 10 per cent accuracy all the properties of pure water found experimentally by observing 10^{22} water molecules for many seconds. Simulating a protein consisting of about a thousand atoms for 10^{-10} seconds in vacuum reveals the way in which these complicated molecules can move. These and other successes suggest that reductionism really does work for molecular systems. As computers keep on getting larger and faster — and there is no end in sight — it should be possible to simulate systems consisting of hundreds of thousands of atoms for times as long as microseconds.

McCammon, who pioneered simulations of protein dynamics in 1976, and Harvey, who applied these methods to nucleic acids six years later, have written the first book in a new field which is rapidly growing into an important branch of science. What is to be expected of such

a book? On the one hand, there is a need for a handbook that provides an in-depth treatment of methods and results using examples to illustrate the limitations of simulation, the approximations behind the potential energy functions, and the importance of thorough analysis. On the other hand, there is a need to summarize and relate the considerable body of results, theoretical, computational and experimental, that is already available. The second path is followed in this rather slim volume, which grew out of two reviews, one by each of the authors. They have not tried to hide the book's origins, and in places it reads more like an extended review than a book.

Chapters 1 to 4 introduce the conceptual foundations and methodology. Although some effort is made to explain the methods, it is not clear that they will be easily understood by a newcomer to the field; also, for the more expert, more emphasis could have been placed on their limitations. Chapter 3, which discusses the dynamics of solvents, proteins and nucleic acids, sometimes presents facts with authority but with no supporting references; it reads more like a concluding discussion than an introduction, and this can be confusing.

Chapter 5, one of the best in the book because of the wealth of results available, considers short-time dynamics, which can be calculated directly by molecular dynamics simulation; there is generally recognized concern that calculations done for the most part in vacuum may not give a true picture of macromolecular dynamics in solution. Chapters 6 and 7 deal with structural changes on a local and global scale, processes that take too long for direct simulation. Instead, use is made of more approximate methods, which are less satisfying than the strictly reductionist approach of molecular dynamics but do allow solvent to be taken into account.

Chapter 8 deals briefly with the dynamics of molecular associations, while the ninth and final chapter surveys future directions. The three appendices, which reproduce mathematical treatments already published in papers, are useful but the formulae (particularly those on pp. 174–175) contain some errors.

In spite of these few misgivings, it must be stressed that this is the first book in a growing field that is of widespread interest. McCammon and Harvey's account is easily read and at the same time rewarding to those seeking deeper insight. With its comprehensive and well-balanced list of references, and its review of all the main developments, it will be an invaluable companion to those interested in the dynamics of biological macromolecules.

Michael Levitt is the M. and B. Levinson Professor in the Department of Chemical Physics, The Weizmann Institute of Science, Rehovot 76100, Israel.

Personal approaches

Peter J. Bowler

Scientific Genius and Creativity: Readings from Scientific American. Introduction by Owen Gingerich. *W.H. Freeman: 1987. Pp.110. Hbk £21.95, \$21.95; pbk £9.95, \$9.95.*

Man Masters Nature: 25 Centuries of Science. Edited by Roy Porter. *BBC Books: 1987. Pp.224. Pbk £4.95.*

BOTH of these books use biographical sketches of famous scientists as a means of informing the general public about the nature and development of science as a whole. Their purpose, format and range of coverage are, however, quite different.

The *Scientific American* reader has the usual large format with plenty of diagrams and illustrations. Its purpose is to throw light on the process of scientific creativity by offering case-studies of great discoveries. To set the scene we have Jacob Bronowski's comparison of scientific and artistic creativity, while a postscript by Gunther Stent explores the meaning of the term 'premature discovery'. The editor's choice of biographical studies was obviously limited by what has already appeared in the pages of *Scientific American*, and although the range extends from Newton to Einstein there is no pretence of offering a complete survey of the history of science.

The real problem for the editor, however, is that he has had to reprint essays which may be up to 30 years old, with only a brief bibliographical note in which to mention more up-to-date studies. I. Bernard Cohen's 1981 account of Newton's work on gravity is still useful, but Loren Eiseley's 1956 mini-biography of Darwin contains only a limited and very outdated outline of the discovery of natural selection. Considering the detailed studies of Darwin's notebooks by a host of scholars in the past couple of decades, it is difficult to see how this piece can be expected to illuminate what we now know to have been an immensely complex act of creative thinking.

The BBC book, by contrast, includes a series of biographies written by specialists well aware of current developments in their fields. The purpose here is to use the biographies as a means of showing how Western science has become so successful — creativity is only one part of the story, along with methodology and the social environment in which the scientist works. There are 17 articles ranging from Aristotle to Watson and Crick, with a fairly balanced coverage of the physical and biological sciences. The problem is that with an average of only a dozen pages each, the authors have to work very hard to give the non-specialist reader an

impression of what their chosen subject achieved and how he did it. Individually the results are surprisingly good, but I am not sure that the man or woman in the street will be able to extract a general message about the nature of modern science from so diverse a collection.

Biography is an obvious vehicle for studying creative thinking, and the *Scientific American* project would have benefited from the BBC's approach of commissioning newly written articles. But I am not persuaded that biography is a suitable format through which to inform the lay public about the general factors responsible for the rise of modern science. A series of biographical studies may have offered a convenient format for the radio programmes upon which the BBC book is based — but is it the format the authors would have chosen to present their ideas to a general audience? Roy Porter's introduction directs his readers to the broad themes explored in the individual articles, but how many of them will realize that the true purpose of the book is to show that science is something more than a series of great discoveries?

Most historians of science now believe that the traditional emphasis on the 'heroes of discovery' has obscured the complexity of the process by which new theories are created and promoted in the scientific community. However hard one tries to dispel the almost myth-like stories surrounding these discoveries, to do so within a biographical format may be to invite misunderstanding. □

Peter J. Bowler is a Lecturer in the Department of History and Philosophy of Science, The Queen's University, Belfast BT7 1NN, UK.



More 'heroes of discovery' — the picture (of Michael Faraday expressing distaste at the polluted state of the Thames) is taken from *The History of Scientific Discovery* edited by Jack Meadows, a similar book to those reviewed here. Publisher is Phaidon, Oxford. In the United States the book is published by Oxford University Press under the title *The Great Scientists*. Price is £19.95, \$35.

Mutual aid society

Gordon H. Orians

Helping and Communal Breeding in Birds: Ecology and Evolution. By Jerram L. Brown. *Princeton University Press: 1987. Pp.354. Hbk \$45, £28.20; pbk \$16.50, £10.40.*

RECENTLY, there has been increasing recognition that genetic contributions to future generations by means of assistance to non-descendant relatives have strongly influenced the evolution of animal societies. The study of social insects provided the impetus for both theoretical and empirical analyses of inclusive fitness, but during the past two decades research on communally breeding birds, based on long-term studies with uniquely marked individuals, has added substantially to our understanding of the evolution of sociality.

This extensive and rapidly growing literature is critically reviewed by Jerram Brown, one of the principal contributors to both theory and field studies. The book is comprehensive in scope. It presents an evaluation of the use of terms (supplemented by an extensive annotated glossary). There are surveys of the geographical and taxonomic distribution of helping and communal breeding, a discussion of theories of inclusive fitness, and detailed treatments of the diverse forms of communal breeding among birds and their theoretical implications.

The book is a storehouse of information, interpretation and clear, if controversial, exposition of complex ideas. Brown makes effective use of a mixture of mathematical formalisms and intuitive arguments, and uses his models to highlight the key inequalities that must obtain if specific types of behaviour are to be favoured by natural selection. Early in the book he establishes the need to separate delayed dispersal, delayed breeding and helping, pointing out that all combinations of these variables exist among birds. He uses this distinction throughout the text to clarify arguments and interpret patterns in areas where failure to keep them apart has caused confusion.

A major contribution of the book is its emphasis on the importance of ecology in the study of evolution of helping and communal breeding. Most sociobiological theories are at least implicitly ecological, but ecological factors are often passed over rather lightly, in part because of a widespread belief that the most important need is to incorporate genetics into sociobiological models. Brown demonstrates that the key to understanding the evolution of communal breeding in many species lies in a better appreciation of how the birds live and how they exploit their environments.

Brown jumps squarely into the middle

of controversies in an area characterized, as is much of science, by a complex mixture of status-seeking and truth-seeking on the part of individual scientists. He does not avoid making hard and sometimes harsh judgements. His purpose is to clarify the issues, but the petulance with which he insists that his own contributions have been unappreciated and misinterpreted is, sadly, likely to stimulate, rather than resolve, future controversies.

The goal of science is always to seek generalizations, but Brown clearly demonstrates that no single hypothesis is adequate to explain the evolution of helping and communal breeding in all species or even the evolution of all components of the social organization of a single species. Evolutionary pathways to sociality have

been diverse. Different species have been influenced by different social and environmental factors.

Comprehensive though the book is, more attention could have been paid to species that appear to fulfil the preconditions for the evolution of helping and communal breeding but do not exhibit such behaviour. We often learn as much from an analysis of what has not happened as we do from studying what has happened. In any case, every student of evolutionary theory, sociobiology or behavioural ecology should read and study this book very carefully. It will be the standard reference on the subject for many years. □

Gordon H. Orians is a Professor in the Department of Zoology, University of Washington, Seattle, Washington 98195, USA.

Models and games

Kendall Preston, Jr

Cellular Automata Machines: A New Environment for Modeling. By Tommaso Toffoli and Norman Margolus. MIT Press: 1987. Pp.259. \$30, £26.95.

FOUR decades ago John von Neumann, the famous mathematician, became interested in the power of digital computation. At that time there were perhaps two fully-configured digital computers in the world. These were not much more than today's programmable hand calculators. Each used a single ALU (Arithmetic Logic Unit), stored a relatively small program and executed instructions slowly.

Von Neumann, however, with his extraordinary vision of the future, began to study the computational capabilities of thousands of ALUs interconnected in a massive planar array where each ALU had direct connections to its neighbours. He called such arrays 'cellular automata'. With von Neumann's premature death, and the impossibility of constructing cellular automata with the technology of the time, cellular automata studies then languished.

In the 1950s, however, Moore and his colleagues at the United States National Bureau of Standards, Unger at Bell Telephone Laboratories, Dinneen and Selfridge at the Massachusetts Institute of Technology (MIT), and Golay and his colleagues at Perkin-Elmer recognized that simulators of cellular automata could be built using fast digital hardware employing the transistor. This was done and, by the 1960s, important research was being carried out with what were called 'cellular logic machines'.

At the same time, work by McCormick and Skolnick at the University of Illinois led to the first true cellular automaton — Illiac IV, a small array (8 × 8) of 64

computing elements. In Britain, International Computers Ltd and University College London produced the first really significant cellular automata, namely the 32 × 32 Distributed Array Processor and the 96 × 96 Cellular Logic Image Processor, respectively, in the late 1970s.

In the 1980s there have been further advances, primarily in the United States. There are now the 128 × 128 Massively Parallel Processor of the National Aeronautical and Space Agency, the 256 × 256 Connection Machine of the Thinking Machines Corp. and the 216 × 384 Geometric Array Parallel Processor of Martin Marietta.

Unfortunately, the book under review does not concern itself with any of the above, except for brief mention of the Connection Machine's capability in running fluid-dynamics computations. Instead, the authors concentrate on the cellular automata simulator CAM-6 which they have developed at MIT. This is not to disparage the book, the contents of which are quite fascinating to the *aficionado* of the cellular automaton, but simply to indicate that the title is a gross misnomer. A better title for the first part of the book would be "Cellular Automata Games for CAM-6" and, for the latter part, "Modeling Physics on CAM-6".

The book has its origins in Fredkin's research at MIT and also in the cellular automata game called 'Life' devised by the mathematician John Horton Conway — both originating in the 1960s. These stimuli led to intensive studies of cellular automata games at MIT, not only using Conway's rule but also many other rules devised by Fredkin's colleagues and students in the Information Mechanics Group. Much of this work is brilliant and it is a pleasure to find it thoroughly documented, with detailed rules and profuse illustrations, in this book. The authors are fascinated by the patterns generated by displaying the contents of the processing elements of the cellular

automaton on the computer screen, both in black and white and in colour. For example, of the display generated by the rule called "TUBE-WORMS", they write: "these animals have plumes that look like delicate flowers, but at the slightest disturbance the 'flower' retracts into a tube and waits a good fraction of minute before coming out again".

Of more practical interest is the latter half of the book which deals with the use of CAM-6 in "modeling of the laws of physics". In this regard the book is not particularly analytical — there is little or no mathematical development — although the heat equation, the diffusion equation and the Navier-Stokes equation do put in an appearance. The illustrations provided do not convince me that the physical models generated by programming CAM-6 are exact. In fact, the authors state that "the gases described by a cellular automaton rule . . . obey the Navier-Stokes equation approximately". The reader is not told how good this approximation is, although we are referred to the literature on hexagonal tessellation cellular automata and are told that *exact* solutions have been obtained. (CAM-6 is not hexagonal but uses the square tessellation.) This, of course, is of great significance and there is a growing interest in applying cellular automata to research on fluid dynamics. General studies of this kind, however, require three-dimensional analyses and, unfortunately, the book mentions work on three-dimensional automata (such as that being carried out at Carnegie-Mellon University) only peripherally.

The colourful pictures are attractive, but otherwise this is not a book for the casual reader. To understand the multitude of CAM-6 rules presented, one must learn the vagaries of the programming language FORTH. An appendix is provided for this purpose, but is in some ways inadequate; the authors' use of the FORTH "case statement", first introduced on page 20, is not properly explained until page 40. But these are minor annoyances to the dedicated reader who is willing to read and re-read to get at the meat of this work.

The book is well referenced with both current and historical material, especially as regards recent applications to physical modelling. Although there are shortcomings, it still deserves serious attention by the rapidly growing community of computer scientists and physicists who now work with cellular automata and cellular automata simulators — not to mention the 'computer artist' who simply wants to sit enthralled watching tube-worms evolve on the computer display. □

Kendall Preston, Jr is a Professor in the Department of Electrical and Computer Engineering, Carnegie-Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213, USA.

The continuing decline of British science

Ben R. Martin, John Irvine, Francis Narin and Chris Sterritt

New and more methodologically advanced publication and citation data for the period up to 1984 indicate that the relative decline of British science is continuing, albeit at a slower rate than in the 1970s.

Two years ago, we published an article charting the decline in British science between 1973 and 1982¹. This was one of the earliest such attempts to assess Britain's scientific performance. The indicators employed, which are based on cross-national publication and citation data, were still in a somewhat experimental state. Among the limitations were the absence of citation data disaggregated to the subfield level, and the fact that the field-level citation figures were derived using an approximate 'probability assignment' method. Other possible technical problems have also been raised by Dr Loet Leydesdorff of Amsterdam University².

In this article, we discuss subsequent methodological improvements to the indicators and the effect they have on the types of bibliometric analysis that can now be undertaken. In particular, we examine what the latest statistics for the period to the end of 1984 reveal about the relative international standing of British science.

In the earlier article, we outlined the reasons why international science indicators are increasingly needed, especially by those concerned with science policy at a national level who have to determine research priorities between as well as within scientific fields. Most importantly, recent years have witnessed the growing economic importance of new technologies such as microelectronics, robotics, new materials and biotechnology, the development of which depends upon the results of basic research³. Industry is now drawing upon academic specialities, funded in the past because of their intrinsic scientific interest, as major inputs to its own longer-term research and development programmes. In other words, areas of previously 'pure' science are being transformed into strategic research. Moreover, as work by CHI Research has shown, this transformation can be comparatively rapid: the time-delay between publication of scientific results and their citation in patent applications is currently as little as two years in certain areas of biology⁴ and electronics⁵. Consequently, policy-makers in mission-orientated government departments and funding agencies require systematic data on national performance across all the basic sciences if they are to make effective decisions as to the fields and specialities in which to concentrate

the limited resources available. Such data are also useful to government more generally to determine whether overall funding is sufficient to maintain a healthy basic research infrastructure (in universities and government institutes) relative to competing industrial nations⁶.

Measuring scientific output

Although various other approaches have been explored, the most comprehensive source of information on national performance in basic science is still the National Science Foundation's *Science Literature Indicators Data-Base*. This is compiled biennially for NSF by CHI Research, and is derived in turn from the *Science Citation Index (SCI)*. The latter is produced by the Institute for Scientific Information (ISI) and currently covers over 3,000 of the world's leading scientific and engineering journals. The CHI/NSF data-base, which now extends from 1973 to 1984, includes over a dozen indicators, although four of these are specific to research in the United States (for example, on active institutions within particular scientific fields). Of the remainder, the most important relate to numbers of publications and citations for some 190 countries and geographical regions, broken down into nine main fields (such as chemistry) and nearly 100 subfields (such as analytical or organic chemistry). One field — psychology — has, however, been excluded here because journals for this area have gradually been transferred from the *SCI* to the *Social Science Citation Index*.

In the data-base, papers are attributed to countries on the basis of the institutional addresses of their authors and to fields and subfields using a journal classification scheme developed by CHI Research. Where a paper is written by researchers from two or more nations, credit is 'fractionated' accordingly between those countries. One problem with the data-base is that there is an inherent bias in the *SCI* towards English-language publications, and for this reason the time-trends for each country are probably more significant than the absolute level of publications and citations. A second problem is that each year ISI continues to increase the number of journals scanned in the *SCI*, but the exact criteria used in

determining which should be added are unclear. The result is that discontinuities are introduced into the time-series, which give rise to certain difficulties with making international comparisons especially at the level of subfields. As a result, CHI has preferred to use a 'constant journal-set' in compiling the various indicators. Because the data-base runs from 1973, most of the indicators are based on the '1973 journal-set' — that is, the 2,000 or so journals scanned by ISI in that year. However, this does mean that papers in new journals established since then are not covered, and newly emerging or fast-growing areas of science may be under-represented. Consequently, CHI has recently begun compiling an alternative data-series based on the 3,000 journals scanned by ISI in 1981. We shall consider data from both journal-sets.

Britain's scientific output

The use of bibliometric data to compare research output in basic science entails making a number of assumptions. One is that most new scientific knowledge is eventually encapsulated in the form of research articles in learned journals. If true, then figures on the relative number of publications, for, say, Britain compared with other countries should give some indication of its changing international scientific standing over time. Table 1 shows national percentage shares of publications (covering all fields of science and engineering combined) for Britain and six other leading scientific countries from 1975 to 1984. For example, of the 261,000 papers published in 1975 (in journals scanned by ISI), US authors produced 37.3%, British authors 9.5% and so on.

Between 1978 and 1981, Japan increased its world share of publications by 13%, while France, the United States and the 'rest of the world' also registered small increases. Britain fell by 2% (compared with a much larger drop for 1975–1978), the Federal Republic of Germany by 6% and the Soviet Union by 8%. Over the three years up to 1984, the British decline has if anything speeded up again slightly to 4%, with France and Germany registering similar falls. Japan's world share continued to grow rapidly, and has probably by now overtaken that of Britain.

Table 1 National percentage shares of scientific publications, 1975–1984

	1973 Journal set				1981 Journal set		'Published-year' data	
	1975	1978	1981	1984	% Change		% Change	
					78–81	81–84	81–84	81–84
Canada	4.3	4.2	4.1	4.3	–2%	+5%	+8%	+8%
France	5.8	5.1	5.2	5.0	+3%	–4%	–4%	–7%
FRG	6.4	6.7	6.3	6.0	–6%	–5%	–5%	–4%
Japan	5.4	6.2	7.1	7.6	+13%	+8%	+7%	+9%
Britain	9.5	8.6	8.4	8.1	–2%	–4%	–2%	0%
United States	37.3	36.7	36.9	36.6	+1%	–1%	–1%	+3%
Soviet Union	7.9	8.2	7.6	7.6	–8%	0%	–1%	–13%
Rest of world	23.4	24.3	24.5	24.9	+1%	+1%	+2%	–1%
Total (k)	261	270	290	286	–	–	–	–
	(100%)	(100%)	(100%)	(100%)				

The above data are drawn from the '1973 journal-set' and may not, therefore, fully reflect research activity in new or fast-growing areas. For this, it is preferable to use the bibliometric data based on the '1981 journal-set' — the penultimate column of Table 1 shows the respective changes in national shares between 1981 and 1984. Comparison of these figures with those for the 1973 journal-set in the previous column reveals that the two data-sets suggest very similar changes — the respective changes in the world publication shares of France, Germany and the United States are identical, while apart from Canada the others differ by only one or two per cent. In other words, the inclusion of an additional 1,000 journals and around 80,000 publications per annum does not markedly affect trends over time in national shares. (It does, however, slightly alter the actual percentage figures for national shares — for example, the 1984 figure for the United States falls from 36.6% to 35.4%, while that for the 'rest of the world' rises from 24.9% to 26.1% and that for the Soviet Union from 7.6% to 7.9%.)

One criticism made of the results presented in our previous paper is that the 1982 figures on national publication shares from the CHI data-base were not fully consistent with those obtained by on-line searching of the *Science Citation Index*. There are several reasons for this, associated with the considerable 'cleaning up' of the ISI data that is required before they can be used to draw reliable international comparisons (for example, correcting misspellings and misattributions of countries in the *SCI*, fractionating collaborative papers, removing letters, obituaries and other non-research articles).

However, another reason is that on-line searching of the *SCI* yields 'published-year' figures, while the CHI publication statistics are 'tape-year' data. To understand the difference, consider the 1984 figures: for the former, these relate to all

publications dated 1984, some of which did not actually become available until 1985 or even later; the latter in contrast are for all publications that entered the *SCI* in 1984, approximately 90% of which would have been dated 1984, with the remainder having earlier dates. On average, one finds that the papers 'lost' as a result of appearing late (because of delays in printing or distribution of journals) are approximately compensated for by those 'gained' from previous years. However, this is not true for the most recent year of the data-base (1984 in this case). The reason is that, although most American and British journals appear promptly, publications from Eastern Europe and the Third World are often delayed. The 1984 'published-year' data consequently understate the relative standing of the Soviet Union and the 'rest of the world', and overstate the positions of Britain and the United States in particular.

This can be seen by comparing the figures in the final column of Table 1 with those in the previous two: according to the 'published-year' data, the Soviet share of publications fell by 13% between 1981 and 1984, while both the 1973 and 1981 journal-set data (based on 'tape year') show little or no change. Conversely, the

'published-year' figures show an apparent improvement in the US position over the three years and no change in the British world share, while both 'tape-year' data-sets show small falls. In short, although one can use 'published-year' statistics for earlier years, for the most recent year in the data-base the 'tape-year' figures must be employed if the problem of delayed journal appearance is to be avoided.

Impact of British science

What has been the impact of British scientific publications relative to those produced in other major industrial countries? To assess this, it is conventional to use citation counts — that is, the number of times published papers are referred to by researchers in subsequent articles. The assumption is that these provide a reasonable indication of the impact of research findings on the international scientific community. In earlier versions of the CHI data-base, rather than assigning citations to papers individually, a 'probability assignment' method was used — for example, if 50% of the papers in *Nature* are by British authors, then 50% of the journal's citations would be credited to Britain. For the most recent version, however, the citation figures are derived by 'exact matching' of citations to papers. Although this technical improvement in the data-base results in some changes in the figures on national shares, the overall international standings and trends are broadly similar.

Figures on national percentage shares of citations for 1976–1984 are given in Table 2. (These include citations to all papers published since 1973, the earliest year covered in the data-base.) It can be seen, for example, that papers published worldwide between 1973 and 1976 earned a total of 1,035,000 citations in 1976, while those appearing between 1973 and 1984 gained 2,492,000 citations in 1984. In 1976, the United States accounted for 51.7% of the total, and by 1984 this had risen slightly to 52.8%. During the same period, the Canadian, French and German shares scarcely changed, while

Table 2 National percentage shares of citations, 1976–1984

	Citing-year data				Cited-year data			
	1976	1980	1984	% Change		1982	% Change	
				76–80	80–84		74–78	78–82
Canada	4.3	4.3	4.2	0%	–2%	3.9	–1%	–6%
France	4.1	4.0	4.0	–2%	0%	4.3	+8%	+7%
FRG	5.4	5.4	5.3	–1%	–1%	6.2	+17%	+5%
Japan	4.1	4.6	5.0	+14%	+7%	6.3	+26%	+27%
Britain	10.9	10.3	9.9	–6%	–4%	8.9	–13%	–8%
United States	51.7	52.3	52.8	+1%	+1%	50.7	–2%	–2%
Soviet Union	2.4	2.0	1.7	–19%	–16%	1.8	–8%	–6%
Rest of world	17.1	17.2	17.2	0%	0%	17.9	+4%	+2%
Total (k)	1,035	2,032	2,492	–	–	707	–	–
	(100%)	(100%)	(100%)			(100%)		

that of the Japanese grew by over 20%. Britain, in contrast, registered a sharp fall of 10%. Again it is important to stress the bias in these data, especially in relation to the Soviet Union, although whether this accounts for the latter's apparent sharp decline is unclear.

The above statistics relate to the impact made in a given year by all previous research (that is, they are 'citing-year' data). CHI also produces citation figures referring to the subsequent impact of research published in a particular year ('cited-year' data), and these are contained in the three right-hand columns of Table 2. For example, all papers published worldwide in 1982 had accumulated a total of 707,000 citations by the end of 1984 (the last year covered in the data-base). These 'cited-year' data indicate an even more rapid movement by Japan (up by 60% between 1974 and 1982), with a much larger fall (of some 20%) evident in the British share compared with eight years earlier. However, they also suggest an upsurge in the relative impact of French and German research, something which has not yet shown up in the 'citing-year' data. The probable explanation for this is that the 'citing-year' figures relate historically to a larger number of years of publications and therefore tend to smooth out changes over time.

So far we have considered national performance in all fields of science and engineering combined. Let us next look at how the seven countries being compared have fared in individual research fields. Table 3 contains statistics for the eight main fields in the CHI/NSF data-base which show how national publication shares changed between 1980 and 1984. Significant features of the table include large falls for France in engineering and biomedical research, and for Germany in biology, clinical medicine and mathematics. Japan increased its share of publications in all eight fields, with particularly rapid growth in clinical medicine, engineering and mathematics. Britain, in

Table 3 Percentage changes in world publication (and citation) shares* by field of research, 1980–1984

	Canada	France	FRG	Japan	Britain	United States	Soviet Union	Rest of world
Biology	+4 (-4)	-5 (-9)	-9 (-1)	+7 (-3)	+1 (-5)	+1 (+3)	+5 (-16)	-4 (+1)
Biomedical research	+14 (+1)	-14 (-4)	-4 (-1)	+17 (+15)	-7 (-5)	+2 (+1)	-1 (-19)	-2 (-2)
Chemistry	-5 (-6)	-6 (+1)	-5 (-3)	+6 (+11)	-11 (-3)	+3 (+2)	+1 (-11)	+1 (-2)
Clinical medicine	+5 (+1)	-6 (0)	-8 (-7)	+23 (+11)	+5 (-2)	-2 (+1)	-7 (-30)	0 (0)
Earth and space sciences	+12 (-7)	-6 (+4)	+1 (+12)	+8 (+3)	-13 (-5)	+4 (+1)	+1 (-5)	-4 (-1)
Engineering and technology	+9 (-7)	-16 (+11)	+2 (-3)	+22 (+17)	-14 (-8)	+3 (-1)	-17 (-10)	0 (+5)
Mathematics	-6 (-4)	+3 (+1)	-19 (-20)	+29 (+2)	-4 (-3)	-8 (+2)	-19 (-2)	+20 (+5)
Physics	-10 (-6)	-2 (+9)	+9 (+11)	+1 (-3)	-4 (-6)	-4 (-2)	0 (-14)	+4 (+8)
All fields combined	+4 (-2)	-8 (0)	-4 (-1)	+12 (+7)	-3 (-4)	0 (+1)	-3 (-16)	0 (0)

* Based on 1973 journal-set. Citations are 'citing-year' data.

contrast, experienced a fall in all fields apart from biology and clinical medicine. Its world citation share also declined in all eight fields (see the figures in parentheses in Table 3). This cannot simply be dismissed as part of a process of other countries 'catching up' with the more mature scientific nations because the United States registered increases in all but two of the fields over the same period.

Table 4 contains more detailed information on trends in Britain's world share of publications and citations broken down by field of research. Although there has been a decline of around an eighth in overall publication share between 1976 and 1984 (from 9.2% to 8.1%), the picture is uneven. Chemistry, engineering and physics have all suffered large decreases since 1976, while in the case of biology and mathematics there has been little change. The only slight encouragement is that the general downward trend seems to have slowed somewhat since 1980, with an

overall decline of 3% between 1980 and 1984 compared with 9% for the preceding four years.

As regards the changing impact of British work on the international research community, the picture is scarcely more optimistic, although there is some indication that the long-term decline which began in the 1970s (and which amounted to a decrease of some 10% over the six years between 1976 and 1982) may have levelled off since 1982. One can see from the two right-hand columns of Table 4 that chemistry, engineering and physics all recorded substantial falls of 12–16% over the period since 1976, although the first two of these fields did show a small increase in world share between 1982 and 1984.

Implications for science-based industry

One finding from our analysis of earlier CHI data was that the weakest parts of British basic science are heavily concentrated in areas with strategic technological importance for the future. This finding is reinforced by the updated statistics reported in Table 5, which lists those subfields in which Britain had a world publication share of more than 10.5% or less than 6% in 1984, along with the corresponding figures for the British citation share (based on 'citing-year' data). One can see that the areas of strength — whether judged in terms of publications or citations — are almost entirely in medical and biological specialities, although geology and electrical/electronic engineering are also quite strong. In contrast, the weaker subfields are generally in industrially related scientific and engineering specialities such as metals and metallurgy, chemical engineering, solid-state physics,

Table 4 Trends in Britain's world publication and citation shares by field of research 1976–1984

	Publications*					Citations†		
	% Change					% Change		
	1976	1980	1984	76–80	80–84	1984	76–80	80–84
Biology	10.5	10.2	10.3	-2%	+1%	12.8	-3%	-5%
Biomedical research	9.2	8.4	7.8	-8%	-7%	10.2	-4%	-5%
Chemistry	8.2	6.8	6.0	-18%	-11%	8.7	-12%	-3%
Clinical medicine	10.1	9.3	9.7	-8%	+5%	10.8	-9%	-2%
Earth and space sciences	8.7	9.5	8.3	+10%	-13%	9.6	+6%	-5%
Engineering and technology	10.9	9.5	8.1	-13%	-14%	10.3	-8%	-8%
Mathematics	6.9	7.0	6.7	+2%	-4%	8.4	+5%	-3%
Physics	6.9	5.9	5.7	-14%	-4%	7.1	-6%	-6%
All fields combined	9.2	8.3	8.1	-9%	-3%	9.9	-6%	-4%

* Based on 1973 journal-set.

† 'Citing-year' data based on 1973 journal-set.

Table 5 Strengths and weaknesses in British science — world publication and citation shares by subfield in 1984

Stronger subfields	% World share		Weaker subfields	% World Share	
	Papers*	Citations†		Papers*	Citations†
Anaesthesiology	21.7	19.0	Metals and metallurgy	3.2	7.9
Respiratory system	16.7	12.8	Pharmacy	3.9	2.1
Parasitology	15.8	21.9	General chemistry	4.1	5.9
Pathology	15.0	15.0	General physics	4.2	5.5
Gastroenterology	14.4	16.8	General mathematics	4.4	6.3
Ecology	14.0	15.8	Nuclear and particle physics	4.5	4.6
Dentistry	13.5	11.0	Nuclear technology	4.9	7.2
General and internal medicine	13.0	16.4	Entomology	5.3	12.9
Geology	11.6	9.1	Chemical engineering	5.4	8.1
Psychiatry	11.3	14.0	Solid-state physics	5.5	6.6
Hygiene and public health	11.0	15.4	General biomedical research	5.6	10.0
Endocrinology	10.8	8.8	Polymers	5.6	6.8
Electrical/electronic engineering	10.8	10.0	Physical chemistry	5.8	8.0
Marine biology and hydrology	10.6	13.5	Cardiovascular system	5.8	6.3
Microbiology	10.5	13.4	Applied physics	6.0	6.3

* Based on 1981 journal-set. Subfields based on less than 1,000 publications per annum have been excluded.

† Citing-year data based on 1973 journal-set.

Table 6 Recent trends in British world publication and citation shares by subfield, 1981–1984

Growing subfields	% Change		Declining subfields	% Change	
	Papers*	Citations†		Papers*	Citations†
Surgery	+43	+8	Biomedical engineering	-32	-13
Applied mathematics	+40	+1	Ophthalmology	-31	-4
Dairy and animal science	+31	+19	Fluids and plasmas	-29	-34
Probability and statistics	+25	+2	Metals and metallurgy	-25	-3
Endocrinology	+22	+1	Obstetrics and gynaecology	-24	+11
Astronomy and astrophysics	+17	+3	Entomology	-23	+3
Paediatrics	+16	-2	Nutrition and dietetics	-20	+3
Anaesthesiology	+13	-4	Chemical engineering	-19	0
Veterinary medicine	+12	-8	Hygiene and public health	-18	-6
			Solid state physics	-17	-3
			Civil engineering	-17	-10
			Mechanical engineering	-16	-8
			Dermatology and venereal disease	-15	-7
			Chemical physics	-15	-4
			General mathematics	-15	-8
			General biomedical research	-14	-3
			Physiology	-13	-9
			Marine biology and hydrology	-13	-1
			Organic chemistry	-13	0
			Materials science	-13	-5
			Analytical chemistry	-12	-6
			Physical chemistry	-12	-7

* Based on 1981 journal-set. Subfields based on less than 1,000 publications per annum have been excluded.

† Citing-year data based on 1973 journal-set.

polymers, physical chemistry and applied physics. What is particularly alarming is the continuing rapid decline in several of these areas: in just two years since 1982, Britain's world publication share in metallurgy dropped from 4.4% to 3.2% and in chemical engineering from 6.2% to 5.4%. That Britain's world citation share in some of these subfields is greater than the corresponding publication figure partly reflects the fact that many of the 1984 citations are to papers published several years earlier when the British publication share was still much higher.

More systematic evidence on recent

trends is contained in Table 6. This lists the subfields in which Britain has either increased or decreased its world publication share by over 12% in the three years since 1981, along with the corresponding figures on changes in citation share. Again, although some subfields such as surgery, applied mathematics, and dairy and animal science have grown appreciably, the overall picture is one of sharp declines in areas of science likely to prove strategically important in the future — for example, biomedical engineering, metallurgy, chemical engineering, solid-state physics, civil and mechanical engineering,

chemical physics, organic chemistry and materials science. Here, the longer-term implications for the chemical and pharmaceutical sector — one of the few science-based industrial areas where technological performance indicators (such as patents) show Britain still has a world-class presence⁷ — are especially worrying.

Conclusions

We have pointed in this article to certain technical limitations in the publication and citation data used previously to assess national performance in science and engineering. However, many of these problems can be considerably reduced by using the greater range of more sophisticated indicators now available — the publication statistics based on the 1981 as well as the 1973 journal-set, citation figures derived by 'exact matching' rather than the 'probability assignment' method, and citation data disaggregated to the subfield level. Yet despite such methodological advances, the overall picture in relation to the comparative international standing of British science remains broadly similar, with the rapid fall in publication and citation world share during the 1970s giving way to a slightly reduced rate of decline in the first half of the 1980s. With companies in both traditional and new industrial sectors becoming increasingly dependent on the results of basic research for maintaining innovative activity and international competitiveness, the state of British science consequently gives great cause for concern. If Britain is to compete more effectively in the international technology market, one important precondition must surely be an increase in the relative strength of its underlying research infrastructure.

The work reported here was funded by a consortium of British government agencies and was primarily carried out at the Technical Change Centre, London. The views expressed in the paper are those of the authors alone. □

1. Irvine, J., Martin, B.R., Peacock, T. and Turner, R. *Nature* **316**, 587–590 (1985).
2. Callon, M. and Leydesdorff, L. *La Recherche* **18**, 412–419 (1987).
3. Irvine, J. and Martin, B.R. *Foresight in Science: Picking the Winners* (Frances Pinter, London, 1984).
4. Narin, F. and Noma, E. *Scientometrics* **7**, 369–381 (1985).
5. Carpenter, M.P., Cooper, M. and Narin, F. *Research Management* **13**, 30–35 (1980).
6. Irvine, J. and Martin, B.R. *Nature* **323**, 591–594 (1986).
7. Patel, P. and Pavitt, K. in *National Institute Economic Review* (November 1987).

Ben R. Martin and John Irvine are Senior Fellows of the Science Policy Research Unit, University of Sussex, Brighton BN1 9RF, UK. Francis Narin is President of CHI Research/Computer Horizons Inc., 10 White Horse Pike, Haddon Heights, New Jersey 08035, USA, and Chris Sterritt is Director of Diamond Systems, 60 Springbank Road, Hesters Way, Cheltenham, Gloucestershire GL51 0NJ, UK. No order of seniority is implied by the author name order. Correspondence should be addressed to John Irvine.

Thermal expansion of sea water associated with global warming

T. M. L. Wigley*[‡] & S. C. B. Raper[†]

* Environmental and Societal Impacts Group, National Center for Atmospheric Research, Boulder, Colorado 80307-3000, USA

† Climatic Research Unit, University of East Anglia, Norwich, NR4 7TJ, UK

The relationship between greenhouse-gas forcing, global mean temperature change and sea-level rise due to thermal expansion of the oceans is investigated using upwelling-diffusion and pure diffusion models. The sensitivities of sea-level to short-timescale forcing and deep-water formation rate changes are examined. The greenhouse-gas-induced thermal expansion contribution to sea-level rise between 1880 and 1985 is estimated at 2–5 cm. Projections are made to the year 2025 for different forcing scenarios. For the period 1985–2025 the estimate of greenhouse-gas-induced warming is 0.6–1.0 °C. The concomitant oceanic thermal expansion would raise sea level by 4–8 cm.

FUTURE increases in the atmospheric concentrations of the greenhouse gases (carbon dioxide, methane, nitrous oxide and chlorofluorocarbons) are expected to result in substantial global-scale warming in future decades. In response to this warming, global mean sea level should change owing to thermal expansion of the oceans and the melting (or accumulation) of land ice^{1–7}. Prediction of these sea-level changes is of importance, because many coastal regions could be adversely affected by even a small sea-level rise. As a prerequisite to such predictions we need to be able to understand past sea-level changes and to predict the future climatic conditions that will affect sea level. Here, we improve on previous estimates of the past and future contributions to sea-level change arising from thermal expansion of the oceans.

Over the past 100 years, while global mean temperature has increased by ~0.5 °C (ref. 8), sea level has risen by 10–15 cm^{2,5,6}. The relative contributions of thermal expansion and ice melting to this sea-level rise are uncertain and estimates vary widely, from a small expansion effect^{4–6} through roughly equal roles for expansion and ice melting^{2,9} to a dominant expansion effect¹⁰.

In principle, modelling the thermal expansion effect would appear to be exceedingly difficult, as a precise determination would require one to be able to model the three-dimensional details of oceanic temperature changes. This is beyond present capabilities; indeed, our knowledge of how deep-ocean temperatures have varied in recent decades and of the physical processes that control any such variations is still rudimentary. In spite of this, simple diffusion or upwelling-diffusion models of the ocean can be expected to give reasonable results for the amount of thermal expansion that might occur in response to greenhouse-gas forcing, even though such models oversimplify oceanic mixing processes. This is because the main contribution to thermal expansion is concentrated in the near-surface layers; this is where both the warming and the thermal expansion coefficient are largest.

To date, only pure diffusion (PD) models have been used to estimate the thermal expansion effect (see, for example, ref. 2), although Revelle³ has included an upwelling term in an approximate way. A pure diffusion model leads to an isothermal steady-state ocean temperature profile. Inclusion of an advective, upwelling term (balanced by high-latitude downwelling) in an upwelling-diffusion (UD) model ensures a realistic steady-state temperature profile. For small times the differences between the

thermal expansion predictions of PD and UD models will be small, provided one begins with a realistic initial profile and both models are calibrated to match past observations. But for times of the order of centuries PD and UD models may give noticeably different results because of the different ways in which they distribute surface heating effects vertically. Here, we use a UD model, calibrated to match past temperature changes within the limits of uncertainty in the model parameters (compare ref. 2), to estimate the thermal expansion effect from 1880 to the present and to predict the range of possible future expansion-related sea-level changes. The results are compared with those obtained using a similarly calibrated PD model.

The model

The model used is a box-upwelling-diffusion energy-balance climate model, an elaboration of the PD model used by Wigley and Schlesinger¹¹. It is similar to the model of Harvey and Schneider¹² and is typical of models currently used to study the transient response of the climate system to continuously changing external forcing¹³. The model differs from that in ref. 11 in that it differentiates land and ocean in both hemispheres, and has upwelling-diffusive oceans. Land-ocean and inter-hemispheric exchange coefficients have been set at values that produce good matches to the seasonal cycles of temperature over land and ocean in response to the annual insolation cycle. The model has an oceanic mixed layer the depth of which is set at a constant 100 m. The results presented here are insensitive to the choice of mixed-layer depth.

The model's output is determined by the imposed forcing, and by internal model parameters which define the rates of land-sea and inter-hemispheric exchange, the strength of ocean mixing processes (mixed-layer depth, diffusivity and upwelling velocity), and the sensitivity of the climate system. The latter is conveniently specified by the equilibrium CO₂-doubling temperature change (ΔT_{2x}), that is, the global mean surface air temperature change which would eventually result if the CO₂ concentration were doubled. The parameters that most affect model output are the diffusivity (κ) and the climate sensitivity. Possible feedbacks involving ocean mixing processes^{12,13} are not considered. We concentrate on a range of κ values (0.5–2.0 cm² s⁻¹) and ΔT_{2x} values (1.5–4.5 °C) which span the limits of current uncertainty^{13–15}. We consider two upwelling cases, a constant upwelling rate of 4 m yr⁻¹ (the standard estimate based on isotope tracer studies^{13,16}) and time-varying upwelling. We also consider the PD case by setting the upwelling rate to zero.

The model computes oceanic thermal expansion using expansion coefficient (β) data from Leyendekkers¹⁷. The

[‡] Permanent address: Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK.

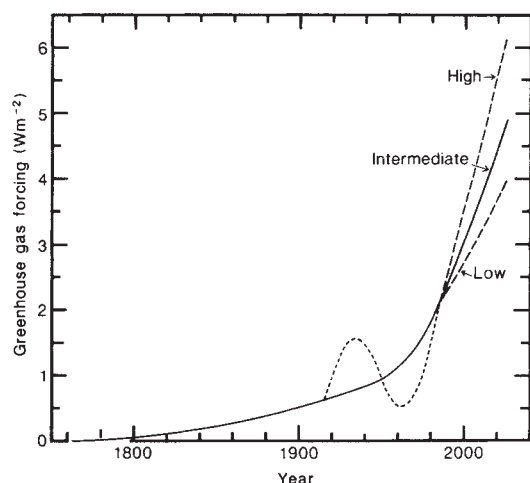


Fig. 1 Assumed forcing based on measured changes in CO_2 , CH_4 , N_2O and CFC concentrations to 1985 and on projections to 2025. The dashed curve shows the superimposed forcing perturbation applied in the Northern Hemisphere for generating the results shown in Fig. 4. The future projections correspond to low, intermediate and high forcing scenarios.

expansion coefficient β ($=\beta(T, p, S)$) where T is temperature, p is pressure and S is salinity) varies widely with temperature and hence with latitude and depth. Vertical variations in β are included explicitly in the model. To account for latitudinal variations we divided each hemisphere into polar, mid-latitude and tropical zones and used the equilibrium mixed-layer results of Manabe and Stouffer¹⁸ to relate the zonal to the hemispheric mean temperature changes (compare ref. 3). (The latitudinal distribution of these changes is similar to that for observed changes over the past century¹⁹.) The thermal expansion results are relatively insensitive to the details of this partitioning. The effect of S on β is small and S is assumed equal to 35‰.

Past temperature and sea-level changes

As temperature changes and expansion are so intimately linked, we must first evaluate the model's simulations of past changes in global mean surface air temperature in relation to the observed warming, namely 0.5°C (ref. 8) between 1880 and 1985 with an uncertainty of at least $\pm 0.1^\circ\text{C}$ ^{20,21}. The forcing changes at the top of the troposphere due to greenhouse-gas concentration changes since 1765 are shown in Fig. 1. Concentration information used comes from refs 22–26 for CO_2 , refs 25 and 27–29 for CH_4 , refs 25 and 27 for N_2O and refs 27 and 30–32 for the CFCs. Concentrations have been converted to radiative forcing using the model results of Ramanathan *et al.*³² and Kiehl and Dickinson³³, with due allowance for overlap effects. Full details are given in ref. 34.

Figure 2 shows the modelled temperature changes from 1880 to 1985, ΔT_0 , for various values of κ and ΔT_{2x} . If greenhouse-gas forcing were the sole mechanism responsible for the 1880–1985 temperature rise, then Fig. 2 would give the range of possible ΔT_{2x} values required for compatibility between model and observations. For an observed warming in the range 0.4 – 0.6°C , the inferred ΔT_{2x} range is 1.2 – 2.2°C , values much lower than recent general circulation model (GCM) estimates which give ΔT_{2x} at $\sim 4^\circ\text{C}$ ^{31,35,36}. However, similarly low values for the climate sensitivity have been obtained in other model-based empirical analyses (for example, 1.6°C in ref. 37).

This apparent discrepancy between observations and recent GCM results can be interpreted in a number of ways: either recent GCM experiments have overestimated the climate sensitivity to a CO_2 change; and/or some additional forcing factor exists which has contributed an overall cooling effect over the past 100 yr, partly offsetting the greenhouse-gas forcing; and/or the upwelling-diffusion parameterization of ocean mixing grossly underestimates the extent of vertical mixing; and/or the 0.4 – 0.6°C estimate of global warming is considerably less than the true warming. GCM uncertainties and/or neglected forcings are likely to be the most important factors. For example, changes in cloud optical properties, which may have a strong negative feedback effect^{38–41} and are not accounted for in current GCMs, could explain at least part of the discrepancy. The

available evidence for additional forcings that are of comparable magnitude to the greenhouse-gas forcing on the century timescale is debatable, but various possibilities have been hypothesized^{37,42–44}.

A useful way to reduce speculation in interpreting Fig. 2, which we will exploit further below, is to assume only that it gives the greenhouse-gas contribution to 1880–1985 temperature changes. If future model results or observations were to show that the 'correct' values for ΔT_{2x} and κ , say, were 3.0°C and $1\text{ cm}^2\text{ s}^{-1}$, then from Fig. 2, the greenhouse-gas contribution to the 1880–1985 warming (namely, ΔT_0) would be 0.78°C . This would then require the existence of a compensating cooling of $0.28 \pm 0.1^\circ\text{C}$ due to other factors.

Figure 3 shows the modelled sea-level change (Δz_0) for the period 1880–1985 due to thermal expansion of the oceans. The range of Δz_0 values compatible with $\Delta T_0 = 0.4$ – 0.6°C is 2.3 – 4.8 cm . This range of values is insensitive to the above-described uncertainties surrounding the greenhouse-gas contribution to the observed warming. We demonstrate this with an example. Suppose that some other external forcing factor (X) operating over the interval 1880–1985 has offset the greenhouse-gas forcing (G) to give a total forcing $T = G - X$, where $X = 0.4G$. With a reduced total forcing compared with G alone, the implied ΔT_{2x}

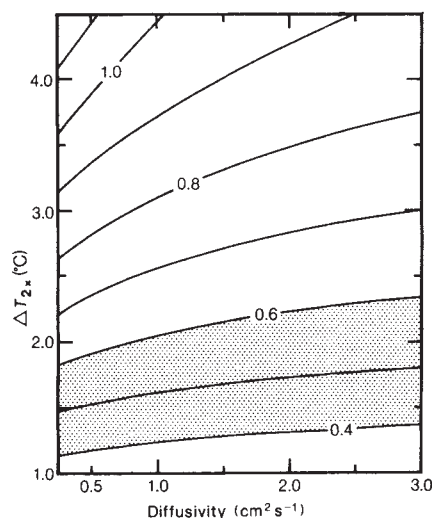
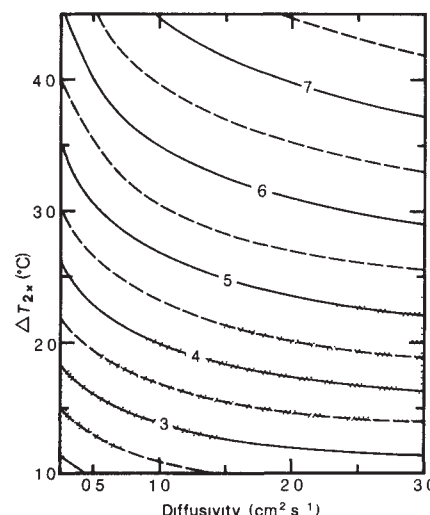


Fig. 2 1880–1985 warming ($^\circ\text{C}$) due to observed increases in greenhouse-gas concentrations for different diffusivities (κ) and equilibrium CO_2 -doubling temperature changes (ΔT_{2x}). For varying κ , the upwelling velocity w has been changed to keep the equilibrium vertical temperature profile unchanged (w/κ kept constant). The results differ very little from those obtained using a constant $w = 4\text{ m yr}^{-1}$. The observed warming range of 0.4 – 0.6°C is shown stippled.

Fig. 3 1880–1985 thermal-expansion-induced sea-level rise (cm) due to observed increases in greenhouse-gas concentrations (Fig. 2) for different diffusivities (κ) and equilibrium CO_2 -doubling temperature changes (ΔT_{2x}). The stippled area gives the range of κ and ΔT_{2x} values compatible with the observed 1880–1985 global warming. For $0.5 \leq \kappa \leq 2.0 \text{ cm}^2 \text{ s}^{-1}$ this implies a contribution of 2–5 cm to sea-level change from thermal expansion.



to give a 0.5°C warming over the period 1880–1985 must be larger—in this case, 3.3°C for $\kappa = 1 \text{ cm}^2 \text{ s}^{-1}$ (as compared with 1.6°C for greenhouse-gas forcing alone). In spite of the very different forcing and the larger implied ΔT_{2x} value, the corresponding Δz is virtually unchanged. For G -minus- X forcing and with the model tuned to give $\Delta T_0 = 0.5^\circ\text{C}$, Δz is 3.44 cm compared with 3.48 cm for the case of G alone. Thus, for century-timescale forcing, Δz is largely determined by the 1880–1985 temperature change, independent of the magnitude of the external forcing which produced this change.

The results shown in Figs 2 and 3 and the link between Δz_0 and ΔT_0 are well approximated by the empirical expression

$$\Delta z_0 = 6.89 \Delta T_0 \kappa^{0.221} \quad (1)$$

for Δz_0 in cm, ΔT_0 in degrees Celsius and κ in $\text{cm}^2 \text{ s}^{-1}$. As Δz_0 and ΔT_0 have approximately the same ΔT_{2x} dependence, the ratio $\Delta z_0 / \Delta T_0$ depends only on κ . A similar result is obtained with a PD model,

$$\Delta z_0(\text{PD}) = 7.98 \Delta T_0 \kappa^{0.315} \quad (2)$$

As the diffusivity value for a PD model fitted to tracer data is higher than for a UD model, PD estimates of Δz_0 may be considerably greater than those obtained with a UD model.

Response to short-timescale forcing

Forcing factors operating on annual to decadal timescales undoubtedly exist and are superimposed on the long-timescale greenhouse effect. These will also affect the sea-level rise at any given time. How significant are these effects? We can say *a priori* that they must be relatively small because the damping effect of oceanic thermal inertia on the temperature response to external forcing increases as the timescale of the forcing decreases. A further evaluation of these effects is, however, of considerable interest because different forcings can produce quite different vertical temperature profile changes (and, hence, sea-level changes) even if the surface temperature changes are the same.

To illustrate this, we consider an extreme example. The largest short-timescale perturbation on the overall global warming trend is the Northern Hemisphere cooling that occurred between ~1940 and 1975. (Note that the warming between 1910 and 1940, which exceeded that expected to result from greenhouse-gas forcing, can be considered as part of this perturbation.) If we simulate these changes in different ways, we can maximize the possible shorter-timescale thermal-expansion-related sea-level fluctuations which might be superimposed on the greenhouse-gas-induced rise of 2.3–4.8 cm. As the temperature perturbations are largest in the Northern Hemisphere, we will use hemispherically specific forcing, as suggested in refs 37 and 45. The vertical ocean diffusivity is assumed to be $1 \text{ cm}^2 \text{ s}^{-1}$, an

acceptable assumption in a sensitivity study like this given the results of Figs 2 and 3.

The two possible causes we consider are an external forcing and a change in the rate of upwelling. Both perturbations are taken to be one-cycle sinusoidal changes spanning the period 1915–85. The variable-upwelling case simulates a change in North Atlantic Deep Water (NADW) formation rate, as the North Atlantic is the main source of deep water in the Northern Hemisphere.

The effects of deep-water formation rate changes on global mean temperatures have been considered previously,^{13,46} but this is the first time that hemispherically specific changes have been considered. (For evidence pointing towards recent NADW changes, see refs 47–49. Major NADW changes are thought to have occurred on the ice-age timescale^{50,51}.) Although NADW changes may well be an important factor in explaining recent temperature fluctuations, especially in the Northern Hemisphere, the main reason for considering them here is because their influence on the vertical ocean temperature profile is radically different from that of an external forcing change.

In both cases the model is calibrated (by varying the climate sensitivity and the amplitude of the forcing or upwelling changes) to produce similar hemispheric and global mean temperature changes. The global mean warming over the period 1880–1985 is set to 0.5°C .

For the external forcing perturbation (shown in Fig. 1), the required amplitude is 0.78 W m^{-2} (that is, a decrease of 1.56 W m^{-2} over 1932.5 to 1967.5; compare this with the greenhouse-gas-forcing increase of 1.81 W m^{-2} over 1880–1985) and the ΔT_{2x} value is 1.77°C . This gives a maximum modelled Northern Hemisphere cooling of 0.20°C , during which the Southern Hemisphere warmed by 0.07°C . Global mean changes are shown in Fig. 4. Both hemispheric and global changes agree well with observations. For the upwelling rate changes, the required amplitude is 0.91 m yr^{-1} (see Fig. 4) and the ΔT_{2x} value is 1.76°C . The corresponding temperature perturbations lead those for the forcing perturbation case by a few years, but they are of the same magnitude—that is, maximum Northern Hemisphere cooling of 0.20°C with concomitant Southern Hemisphere warming of 0.07°C . Global mean changes are shown in Fig. 4.

The sea-level effects of these two perturbations are quite different (see Fig. 4), even though both produce similar surface temperature effects. For the two cases, the vertical profiles of the temperature changes agree only at the surface. Clearly, for any given surface temperature history, there is no unique history for sub-surface temperature changes, and hence no unique sea-level time series. In our analysis, however, this is a secondary effect; both perturbations are small relative to the century-timescale influence of the greenhouse effect—deviations of

<0.7 cm compared with a total change of 3.4 cm. The implied uncertainty in the overall rise to 1985 is therefore $\sim \pm 20\%$, owing to short-timescale forcing uncertainties.

The 2.3–4.8 cm range of values for Δz obtained here is compatible with the previous estimates of Gornitz *et al.*². These authors were the first to examine systematically the thermal expansion effect using a model-based approach; a number of other estimates (for instance, refs 7 and 9) are based on different interpretations of Gornitz *et al.* They used a PD model from ref. 42 which considered the world's oceans as a single column. They applied CO_2 , volcanic and solar forcing and tuned ΔT_{2x} and κ to match model output with the Hansen *et al.*⁴² estimate of global mean temperature changes since 1880. Because they ignored the forcing of other greenhouse gases, they were able to obtain a match for much higher ΔT_{2x} values than obtained here. Also, because their forcing history contains shorter-timescale fluctuations than we have used, their modelled fluctuations in sea level show considerable short-timescale variability. In spite of these differences, they obtain similar thermal expansion estimates to ours, that is, 2.3 cm over the period 1880–1980 and 4.1 cm over the period 1900–80 (for $\Delta T_{2x} = 2.8^\circ\text{C}$ and $\kappa = 1.2\text{ cm}^2\text{ s}^{-1}$). Further support for the values calculated here comes from the empirical estimates of Barnett^{5,6}, who judges the 1880–1980 expansion effect to be <5 cm.

Future thermal expansion effects

We now return to the constant-upwelling-rate case and project the thermal expansion calculations to the year 2025. The many uncertainties involved in making such a projection may be grouped in two categories, those associated with the projected forcing and those associated with modelling the implied thermal expansion. To cover forcing uncertainties, we have used low, intermediate and high values for the forcing (Fig. 1). The intermediate case uses the best estimates of future greenhouse-gas concentrations from ref. 32, but with due consideration given to other estimates in the literature. Concentrations assumed for the year 2025 are (with 1985 values given in brackets): CO_2 , 436 p.p.m.v. (346 p.p.m.v.); CH_4 , 2,235 p.p.b.v. (1,643 p.p.b.v.); N_2O , 366 p.p.b.v. (309 p.p.b.v.); F11, 0.93 p.p.b.v. (0.22 p.p.b.v.); F12, 1.59 p.p.b.v. (0.37 p.p.b.v.). (Many other CFCs have been included in pre-1985 and future forcing estimates; F11 and F12 contribute over three-quarters of the total CFC forcing.) Full details of this intermediate-forcing scenario are given in ref. 34. For the low-forcing scenario we have used post-1985 forcing changes that are two-thirds of those used in the intermediate case; for the high-forcing case the post-1985 forcing changes have been multiplied by 1.5. The corresponding projections agree well with other estimates in the literature.

It is somewhat more difficult to account for model uncertainties. We do this by using some remarkable, robust features of the model projections. First, we use the fact that the temperature change ratio, $\Delta T_1/\Delta T_0$ (where subscript 0 refers to 1880–1985 and 1 refers to 1985–2025), is practically independent of the assumed values of ΔT_{2x} and κ ⁵². $\Delta T_1/\Delta T_0$ depends only on the 1985–2025 to 1880–1985 forcing ratio, $\Delta F_1/\Delta F_0$. The relationship is nearly linear over a wide range of values of ΔF_1 . (For the three models here, $\Delta F_1 = 1.77$, 2.65 and 3.97 W m^{-2} , while $\Delta F_0 = 1.81\text{ W m}^{-2}$.) For $\Delta T_{2x} = 1.5$ – 4.5°C and $\kappa = 0.5$ – $2.0\text{ cm}^2\text{ s}^{-1}$, we can estimate ΔT_1 to within a few per cent using

$$\Delta T_1 = (0.308 + 0.49\Delta F_1)\Delta T_0 \quad (3)$$

A similar relationship results if one uses a PD rather than a UD model; indeed, equation (3) gives results which are accurate for a PD model to within a few per cent. As an example, equation (3) implies that, if the greenhouse-gas contribution to ΔT (1880–1985) were 0.6°C , then (for the intermediate-forcing case) the greenhouse-gas contribution to ΔT (1985–2025) would be $\sim 1.6 \times 0.6 = 0.96^\circ\text{C}$, independent of ΔT_{2x} , κ and the model structure.

Next we use the fact that the ratio $\Delta z_1/\Delta T_1$ is virtually

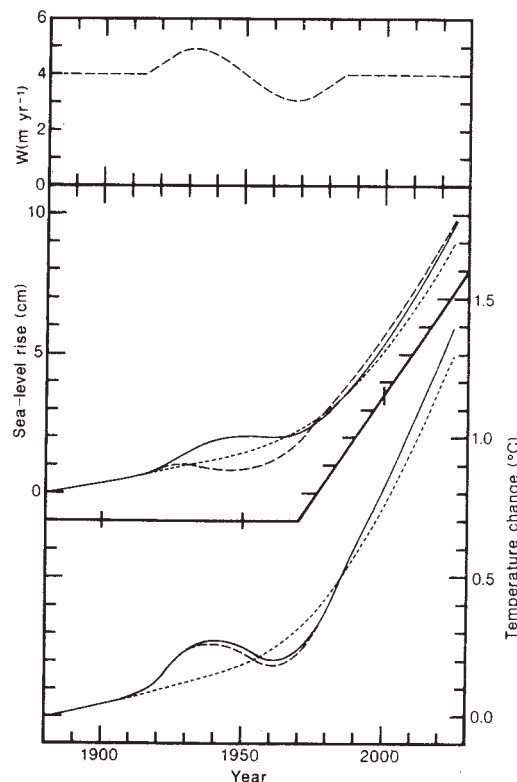


Fig. 4 Global mean temperature and sea-level changes for two different explanations of the mid-twentieth-century cooling of the Northern Hemisphere and the disparate Northern and Southern Hemisphere temperature trends. Dotted curves, greenhouse-gas forcing alone; dashed curves, greenhouse-gas forcing coupled with upwelling velocity changes in the Northern Hemisphere as shown in the top panel; full curves, greenhouse-gas forcing with additional sinusoidal forcing in the Northern Hemisphere only, as shown in Fig. 1. In all three cases the model has been calibrated to produce a global warming of 0.5°C between 1880 and 1985.

independent of ΔT_{2x} and depends only weakly on κ . $\Delta z_1/\Delta T_1$ does, however, depend on the chosen forcing scenario. This result can be expressed in the form (compare equation (1))

$$\Delta z_1 = f(\Delta F_1)\Delta T_1\kappa^{0.221} \quad (4)$$

Equations (3) and (4) can be combined to give

$$\Delta z_1 = (4.13 + 2.65\Delta F_1)\Delta T_0\kappa^{0.221} \quad (5)$$

Equation (5) describes the model results to better than $\pm 4\%$ for a wide range of ΔT_{2x} and κ values (ΔT_{2x} is included implicitly through the term ΔT_0). Equation (5) also implies that Δz_1 is non-zero even if $\Delta F_1 = 0$ (for this extreme case, equation (5) has a maximum error of $\sim 20\%$), a result which is a necessary consequence of the oceanic lag effect and the current disequilibrium between global mean temperature (and sea level) and the greenhouse-gas forcing.

An expression similar to equation (5) may be derived for PD model results:

$$\Delta z_1(\text{PD}) = 1.14(4.13 + 2.65\Delta F_1)\Delta T_0\kappa^{0.301} \quad (6)$$

Thus, PD estimates of future expansion may noticeably exceed UD estimates (for $\kappa > 0.2$).

Consider an example based on the UD result, equation (5). Suppose that the greenhouse-gas contribution to the 1880–1985 global warming is 0.4 – 0.6°C , and that the future forcing follows the intermediate scenario. The implied 1985–2025 greenhouse-gas-induced warming is 0.64 – 0.96°C , and the corresponding range for Δz_1 , allowing for uncertainties in κ in the range 0.5 – $2.0\text{ cm}^2\text{ s}^{-1}$, is 3.8 – 7.8 cm . We consider this range of values to be the best estimate of future oceanic thermal expansion effects over the 1985–2025 interval. Details of changes as a

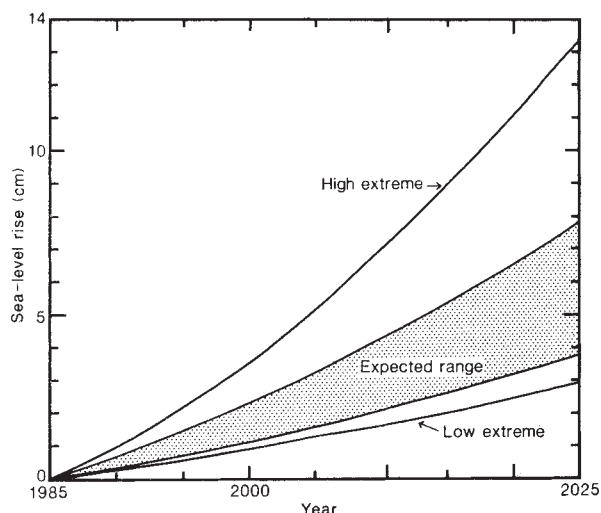


Fig. 5 Thermal-expansion-related sea-level changes over the period 1985–2025. The lower and upper curves correspond to estimated extreme limits (see text). The stippled area gives the range of most likely future changes.

function of time between 1985 and 2025 are shown in Fig. 5.

The lower value is probably close to a lower bound, as it is unlikely that the greenhouse-gas contribution to $\Delta T(1880-1985)$ is $<0.4^\circ\text{C}$ (0.4°C would require ΔT_{2x} to be $1.2-1.3^\circ\text{C}$). For the low-forcing case and $\Delta T_0 = 0.4^\circ\text{C}$, Δz_1 is 2.9 cm (equation (5) gives 3.0 cm). To obtain an upper bound, consider first the accepted upper bound for ΔT_{2x} of 4.5°C . With $0.5 \leq \kappa \leq 2.0 \text{ cm}^2 \text{ s}^{-1}$, this would imply a greenhouse-gas contribution to 1880–1985 temperature changes of $\Delta T_0 = 0.93-1.09^\circ\text{C}$ (see Fig. 2). Such large values would require a very large additional (negative) forcing to be operating on the century timescale to be compatible with observations. A more realistic upper bound to ΔT_0 is 0.8°C (implying ΔT_{2x} in the range $2.8-3.5^\circ\text{C}$). For the high-forcing case, the corresponding value for ΔT_1 is 1.80°C and the maximum ($\kappa = 2.0 \text{ cm}^2 \text{ s}^{-1}$) value of Δz_1 is 13.4 cm (equation (5) gives 13.7 cm). Figure 5 shows the time evolutions of these upper and lower extremes.

The Δz_1 range of 3.8–7.8 cm, with extreme limits of 2.9–13.4 cm, is noticeably less than other estimates of the future

thermal expansion effect. Gornitz *et al.*² give a value of 20 cm for the change from 1980 to 2050 which can be translated to a 1985–2025 change of ~ 10 cm. Hoffman *et al.*⁹ give values of 6.5, 13.1 and 18.3 cm as low, intermediate and high values for 1980–2025 (equivalent to 6.0, 12.5 and 17.6 cm for 1985–2025). More recently these authors give lower and upper bounds of 5.7–10.8 cm (1985–2025)⁵³. These are the only other publications where time-dependent, transient-response estimates have been made which can be compared with the present work. The differences between these estimates and the present work arise from model differences (all other work has been based on PD models which produce larger expansion effects; see equations (2) and (6)), from differences in the assumed future forcing scenarios, and from differences in the way spatial variations in β have been accounted for and model parameter ranges have been chosen.

Conclusions

These results imply only a small thermal expansion contribution to past sea-level changes (2–5 cm compared with the estimated observed rise of 10–15 cm over the period 1880–1980). The contribution of the melting of small glaciers over 1900–61 has been estimated at $0.46 \pm 0.26 \text{ mm yr}^{-1}$ (ref. 54), that is, 2–7 cm over the 1880–1980 period. Long-timescale isostatic rebound effects could account for a further 2 cm (ref. 2). These figures could imply a substantial additional contribution to sea-level rise from melting of the large ice sheets in Antarctica and Greenland, but they are also compatible with a negligible contribution from this source. Our projections of greenhouse-gas-induced sea-level rise due to thermal expansion between 1985 and 2025 are also relatively small, 4–8 cm, accompanied by a global mean warming in the range $0.6-1.0^\circ\text{C}$. Estimating future changes in sea level therefore depends crucially on predicting the future melting of land-based glaciers and ice sheets, a daunting task.

This work was supported by grants from the US Department of Energy, Carbon Dioxide Research Division and the European Economic Community Climate Programme. The National Center for Atmospheric Research is sponsored by the NSF. Comments from R. A. Warrick helped to improve this paper substantially. Radiative transfer model results were kindly provided by J. T. Kiehl.

Received 26 May; accepted 17 September 1987.

1. Etkins, R. & Epstein, E. S. *Science* **215**, 287 (1982).
2. Gornitz, V. L., Lebedeff, L. & Hansen, J. *Science* **215**, 1611 (1982).
3. Revelle, R. in *Changing Climate*, Carbon Dioxide Assessment Committee, National Research Council, 433–448 (National Academy Press, Washington DC, 1982).
4. Polar Research Board, NRC (Meier, M. F. *et al.*) *Glaciers, Ice Sheets and Sea Level: Effect of a CO₂-Induced Climatic Change* Report ER 60235-1 (US Department of Energy, Washington DC, 1985).
5. Barnett, T. P. *Mon. Weath. Rev.* **112**, 303 (1984).
6. Barnett, T. P. in *Detecting the Climatic Effects of Increasing Carbon Dioxide* (eds McCracken, M. C. & Luther, F. M.) 91–107 (US Department of Energy, Washington DC, 1985).
7. Robin, G. deQ. in *Greenhouse Gases, Climatic Change, and Ecosystems* (eds Bolin, B., Döös, B. R., Jäger, J. & Warrick, R. A.) 323–359 (Wiley, Chichester, 1986).
8. Jones, P. D., Wigley, T. M. L. & Wright, P. B. *Nature* **322**, 430 (1986).
9. Hoffman, J. S., Keyes, D. & Titus, J. G. *Projecting Future Sea-Level Rise*. Report PM-221 (US Environmental Protection Agency, Washington DC, 1983).
10. Robock, A. *Science* **219**, 996 (1983).
11. Wigley, T. M. L. & Schlesinger, M. E. *Nature* **315**, 649 (1985).
12. Harvey, L. D. D. & Schneider, S. H. *J. geophys. Res.* **90**, 2207 (1985).
13. Hoffert, M. I. & Flannery, B. P. in *Projecting the Climatic Effects of Increasing Carbon Dioxide* (eds McCracken, M. C. & Luther, F. M.) 149–190 (US Department of Energy, Washington DC, 1985).
14. McCracken, M. C. & Luther, F. M. (eds) *Projecting the Climatic Effects of Increasing Carbon Dioxide* Report ER-0237 (US Department of Energy, Washington DC, 1985).
15. Bolin, B., Döös, B. R., Jäger, J. & Warrick, R. A. (eds) *Greenhouse Gases, Climatic Change, and Ecosystems* (Wiley, Chichester, 1986).
16. Broecker, W. S. & Peng, T. H. *Tracers in the Sea* (Lamont-Doherty Geological Observatory, Columbia University, New York, 1982).
17. Leyendekkers, J. V. *Thermodynamics of Seawater Part 1* (Dekker, New York, 1976).
18. Manabe, S. & Stouffer, R. J. *J. geophys. Res.* **85**, 5529 (1980).
19. Wigley, T. M. L. & Jones, P. D. *Nature* **292**, 205 (1981).
20. Wigley, T. M. L., Angell, J. K. & Jones, P. D. in *Detecting the Climatic Effects of Increasing Carbon Dioxide* (eds McCracken, M. C. & Luther, F. M.) 55–90 (US Department of Energy, Washington DC, 1985).
21. Wigley, T. M. L., Jones, P. D. & Kelly, P. M. in *Greenhouse Gases, Climatic Change, and Ecosystems* (eds Bolin, B., Döös, B. R., Jäger, J. & Warrick, R. A.) 271–322 (Wiley, Chichester, 1986).

22. Friedli, H., Löttscher, H., Oeschger, H., Siegenthaler, U. & Stauffer, B. *Nature* **324**, 237 (1986).
23. Gammon, R. H., Sundquist, E. T. & Fraser, P. J. in *Atmospheric Carbon Dioxide and the Global Carbon Cycle* (ed. Trabalka, J.) 25–62 (US Department of Energy, Washington DC, 1985).
24. Neftel, A., Moor, E., Oeschger, H. & Stauffer, B. *Nature* **315**, 45 (1985).
25. Pearman, G. E., Etheridge, D., de Silva, F. & Fraser, P. J. *Nature* **320**, 248 (1986).
26. Siegenthaler, U. & Oeschger, H. *Tellus* **39B**, 140 (1987).
27. Rasmussen, R. A. & Khalil, M. A. K. *Science* **232**, 1623 (1986).
28. Rinsland, C. P., Levine, J. S. & Miles, T. *Nature* **318**, 245 (1985).
29. Stauffer, B., Fischer, G., Neftel, A. & Oeschger, H. *Science* **229**, 1386 (1985).
30. Dickinson, R. E. & Cicerone, R. J. *Nature* **319**, 109 (1986).
31. Hansen, J. *et al.* in *Climate Processes and Climate Sensitivity* (eds Hansen, J. & Takahashi, T.) 130–163 (Am. Geophys. Union, Washington DC, 1984).
32. Ramanathan, V., Cicerone, R. J., Singh, H. B. & Kiehl, J. T. *J. geophys. Res.* **90**, 5547 (1985).
33. Kiehl, J. T. & Dickinson, R. E. *J. geophys. Res.* **92**, 2991 (1987).
34. Wigley, T. M. L. *Climate Monitor* **16**, 14 (1987).
35. Washington, W. M. & Meehl, G. A. *J. geophys. Res.* **89**, 9475 (1984).
36. Wetherald, R. T. & Manabe, S. *Clim. Change* **8**, 5 (1986).
37. Gilliland, R. L. & Schneider, S. H. *Nature* **310**, 38 (1984).
38. Charlcock, T. P. *J. atmos. Sci.* **38**, 661 (1981).
39. Charlcock, T. P. *Tellus* **34**, 245 (1982).
40. Somerville, R. C. J. & Remer, L. A. *J. geophys. Res.* **89**, 9668 (1984).
41. Twomey, S. A., Piepgrass, M. & Wolfe, T. L. *Tellus* **36**, 356 (1984).
42. Hansen, J. *et al.* *Science* **213**, 957 (1981).
43. Gilliland, R. L. *Clim. Change* **4**, 111 (1982).
44. Viniklov, K. Ya. & Groisman, P. Ya. *Meteorologiya i Gidrologiya* **1981**(11), 30 (1981).
45. Jones, P. D., Raper, S. C. B. & Wigley, T. M. L. *J. Clim. appl. Meteor.* **25**, 1213 (1986).
46. Watts, R. G. *J. geophys. Res.* **90**, 8067 (1985).
47. Brewer, P. G. *et al.* *Science* **222**, 1237 (1983).
48. Swift, J. H. in *Climate Processes and Climate Sensitivity* (eds Hansen, J. & Takahashi, T.) 39–47 (Am. Geophys. Union, Washington DC, 1984).
49. Roemmich, D. & Wunsch, C. *Nature* **307**, 447 (1984).
50. Boyle, E. A. & Keigwin, L. D. *Science* **218**, 784 (1982).
51. Broecker, W. S., Peteet, D. M. & Rind, D. *Nature* **315**, 21 (1985).
52. Wigley, T. M. L. *Geophys. Res. Lett.* (in press).
53. Hoffman, J. S., Wells, J. & Titus, J. G. in *Iceland Coastal and River Symposium* (ed. Sigbjarnason, G.) 245–266 (National Energy Authority, Reykjavik, 1986).
54. Meier, M. F. *Science* **226**, 1418 (1984).

cDNA sequence of human apolipoprotein(a) is homologous to plasminogen

John W. McLean, James E. Tomlinson, Wun-Jing Kuang*, Dan L. Eaton, Ellson Y. Chen*, Gunther M. Fless†, Angelo M. Scanu† & Richard M. Lawn

Departments of Cardiovascular Research and * Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

† Department of Medicine, Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, USA

Lipoprotein(a) is an LDL-like lipoprotein whose concentration in plasma is correlated with atherosclerosis. The characteristic protein component of lipoprotein(a) is apolipoprotein(a) which is disulphide-linked to apolipoprotein B-100. Sequencing of cloned human apolipoprotein(a) complementary DNA shows that it is very similar to human plasminogen. It contains a serine protease domain and two types of plasminogen-like kringle domains, one of which is present in 37 copies.

LIPOPROTEIN(a) concentration in human plasma is significantly correlated with the risk of heart disease and was first described by Berg¹ as an antigen present in about 30% of a Northern European population. It is now apparent that essentially all plasmas contain this lipoprotein, Lp(a), in concentrations that range from less than 1 to more than 100 mg dl⁻¹ (refs 2, 3). Recent studies demonstrate the correlation between high plasma concentrations of Lp(a) and atherosclerosis, assessed by myocardial infarction, angiography or ultrasonic imaging of arterial blood flow⁴⁻⁹.

Lp(a) closely resembles low density lipoprotein (LDL) in lipid composition and the presence of apolipoprotein B-100 ($M_R \sim 550,000$, 550K), the ligand by which LDL binds to the LDL receptor¹⁰. The distinguishing feature of Lp(a) is the presence of apolipoprotein(a) bound to apoB-100 by a disulphide linkage¹¹⁻¹³. Apolipoprotein(a) itself is a large glycoprotein that exhibits size heterogeneity among individuals with isoforms that migrate on SDS-PAGE in the range of $M_R \sim 300,000$ –700,000¹¹⁻¹⁶. A plausible cause for this heterogeneity of size will be discussed.

Because in other aspects it resembles LDL, the postulated atherogenicity of Lp(a) is probably due to the presence of the apolipoprotein(a), termed apo(a), component. We have therefore undertaken the structural analysis of apo(a). In this manuscript, we present the sequence of cloned human apo(a) cDNA, which shows extreme similarity to plasminogen, as well as a remarkable degree of internal repetition. Based on the sequence data, we suggest a model for the evolution of the apo(a) gene and possible functions of this unusual apolipoprotein.

Isolation and sequence of cDNA clones

Apo(a) was purified from a single donor, and the intact protein and isolated peptides derived from limited proteolysis were both subjected to amino-terminal sequence analysis¹⁷. Marked sequence similarity to plasminogen was noted. Plasminogen is a plasma serine protease zymogen that consists of five homologous and tandemly repeated domains called kringles, and a trypsin-like protease domain; the kringles share about 50% pairwise amino acid identity. We synthesized a 75-nucleotide probe, designated PL.1, which duplicated the DNA sequence encoding part of plasminogen kringle 4 (ref. 18), a region noted to be similar to apo(a). A cDNA library derived from hepatic RNAs of five different persons (whose Lp(a) levels were unknown) was constructed and screened with the PL.1 probe. Of one million plaques screened, 18 hybridizing clones were recovered. The sequence of one clone, designated λ a1, agreed with the characteristics of apo(a) predicted by our protein sequence data. The cDNA insert in clone λ a1 commences with the sequence CAAAATG. The ATG begins an open reading

frame extending throughout the clone (Fig. 1). The initial 19 predicted amino acids resemble a typical hydrophobic signal prepeptide. Thereafter begins a sequence which agrees in 14 of 15 residues with our amino terminal sequence of purified apo(a) (ref. 17). The only discrepancy is that position 3 is encoded as serine in cDNA, but was erroneously assigned as proline in protein sequence analysis. Clone λ a1 represented the least ambiguous apo(a) clone we initially characterized because, despite the extreme similarity of apo(a) to plasminogen, a distinguishable sequence feature was available in that the amino-terminus of apo(a) is unlike the amino-terminus of plasminogen, but is similar to amino acid 350 at the beginning of plasminogen kringle 4 (Fig. 2). This is precisely the structure of clone λ a1. In addition, protein sequencing implied the existence of several closely-related kringle 4-like sequences¹⁷. Two of these occur in λ a1, where they agree with the sequence of the PL.1 probe by 62 and 61 out of 75 nucleotides, respectively.

All subsequent libraries were constructed from a single human liver RNA sample which was chosen on the basis of dot blot hybridization of the five available liver RNA samples with λ a1. Libraries were screened with λ a1, and subsequently with λ a41. Additional clones representing the 5' end of apo(a) messenger RNA were identified by rescreening positive plaques with a synthetic 30-base oligonucleotide that spanned the breakpoint of apo(a) and plasminogen similarity in the signal peptide region (Fig. 1b, Fig. 2).

The 5.2-kilobase (kb) clone λ a41 represents the 3' portion of apo(a) mRNA. λ a41 contains homology to plasminogen 3' untranslated sequences, regions encoding its kringle 5 and protease domains, and 12 tandem copies of the plasminogen kringle 4-like sequence. Figure 1a shows the location of independent apo(a) cDNA clones that were sequenced.

To compare apo(a) and plasminogen cDNA sequences, we also sequenced plasminogen clones obtained from our initial screen (J. T., data not shown) because at this time only a partial sequence had been published.

Northern blot hybridizations were performed to determine the overall size of apo(a) mRNA, to obtain an estimate of the number of kringle 4-like repeat units, and to confirm the lack of the kringle 1–3 region of plasminogen. Lanes 1 and 2 of Fig. 3 show human liver mRNA and HepG2 hepatocarcinoma cell mRNA probed with labelled apo(a) clones. The uppermost hybridizing band, measuring about 14 kb, was identified as apo(a) mRNA. Cross-hybridization between apo(a) and plasminogen sequences is expected. The most intense band of 2.9 kb represents plasminogen mRNA, consistent with the size of plasminogen cDNA. To help interpret the Northern blot, a 'plasminogen-specific' restriction fragment probe (encoding kringles 1–3) from plasminogen cDNA was hybridized to the same filter

after ^{32}P signal decay and high temperature probe stripping. Lanes 3 and 4 show the filter from lanes 1 and 2 reprobbed with this fragment. The 2.9-kb band again hybridizes strongly, as do bands corresponding to 18S and 28S ribosomal RNA. The bands measuring 6.6 and 3.6 kb also rehybridized. These could represent incompletely spliced plasminogen RNAs or other kringle 1–3 containing mRNAs. Significantly, the 14-kb band did not hybridize with this probe (even with overexposure), demonstrating that apo(a) mRNA from HepG2 cells and liver is unlikely to contain close homology to the kringle 1–3 encoding region of plasminogen. Further evidence that this portion of the plasminogen gene is not contained within apo(a) is that no kringle 1, 2 or 3-like sequences were revealed either by sequence analysis of 14 apo(a) cDNA clones, by hybridization analyses of many other clones, or by partial peptide sequence of ~30 proteolysis-derived fragments. To obtain a more accurate size estimate of the large apo(a) mRNA, the filter was probed a third time with a 5-kb fragment of a human apoB-100 cDNA clone (generously provided by M. Reuben). From cDNA sequence analysis, apoB-100 mRNA measures 14 kb (refs 21, 22). Lane 5 shows the previous lane 4, reprobbed with apoB without stripping the kringle 1–3 probe. The only new band to appear lies close to the position of apo(a) mRNA. As we shall discuss later, polymorphism in the size of apo(a) mRNA might be expected.

Overall structure of apo(a)

Figure 4 compares the general features of the apo(a) and plasminogen mRNA sequence and the predicted protein sequences. Apo(a) looks like a plasminogen gene gone awry. The plasminogen mRNA is ~2.9 kb and encodes a 19 residue prepeptide and a mature protein of 791 amino acids (glycoprotein $M_R \sim 92,000$). (cDNA sequencing (ref. 19 and J. T., unpublished data) indicate that plasminogen is one amino acid longer than previously proposed by peptide sequencing, owing to a third consecutive Ile at position 67.) Apo(a) mRNA encodes a predicted mature protein of 4,529 amino acids (translated $M_R \sim 500,000$) with a 19-residue presequence. At the 5' end, the two cDNAs share near identity of 93 base pairs (bp) which comprise 45 bases of 5' untranslated sequence and 16 codons of signal peptide sequence. Thereafter the two sequences abruptly diverge; local DNA sequence identity falls from 100% to 30% (Fig. 2). The apo(a) sequence continues with three more residues of signal sequence and the mature amino-terminus that is similar to kringle 4 of plasminogen, whereas plasminogen cDNA continues to encode its amino-terminal 'tail' region and kringles 1, 2 and 3. It is noteworthy that this relative deletion point coincides with the location of an intron in the plasminogen gene that separates its kringles 3 and 4¹⁸. The beginning of the exon which now encodes the first part of plasminogen kringle 4 encodes in apo(a) the final 3 residues of its signal peptide and the mature amino terminus. Thereafter follows a 37-fold multiplication of the exons encoding the kringle 4-like domain, and the non-repeated region of apo(a) homologous to plasminogen kringle 5, the serine protease domain, and the 3' untranslated region.

The degree of intragenic homology in apo(a) is striking. Of the 37 repeats of the 342-bp kringle 4-like unit, 24 are identical in nucleotide sequence; 4 more are identical, with a sequence that differs in only 3 nucleotides from the 24-fold repeat; and the remaining copies differ by 11 to 71 nucleotides. Kringle 33 has a relative deletion of 24 bp. The number of identical repeats is based on the size of mRNA, and is probably 24 ± 2 . The most 3' of the repeats contains the greatest similarity to plasminogen kringle 4 sequence. The presence of the repeated sequences affected strategies for obtaining overlapping cDNA clones and for DNA sequencing. Although recombination during bacterial propagation was evident at times, we do not believe that the presence of identical repeat sequences is an artefact of cloning for the following reasons: kringle-encoding regions with 100% sequence identity have been sequenced from 10 independent cDNA clones and are thus not a product of expansion due to

recombination within a given plasmid or bacteriophage. Second, genomic Southern blots with various restriction enzymes show hybridizing bands of ~30-fold higher intensity than single copy hybridizing bands (data not shown). The size and restriction sites within the multi-copy genomic fragment are consistent with information derived from sequenced intron-containing clones (manuscript in preparation). Hence, uncloned genomic DNA appears to contain multiple repeats of the kringle 4-like region of apo(a) which have extreme conservation of both intron and exon sequences. Finally, apo(a) clones were obtained from a recombination deficient cloning system (λ -ZAP vector and XL1-Blue recA1 bacteria; Stratagene). Sequence analysis of these clones were consistent with previously characterized cDNAs.

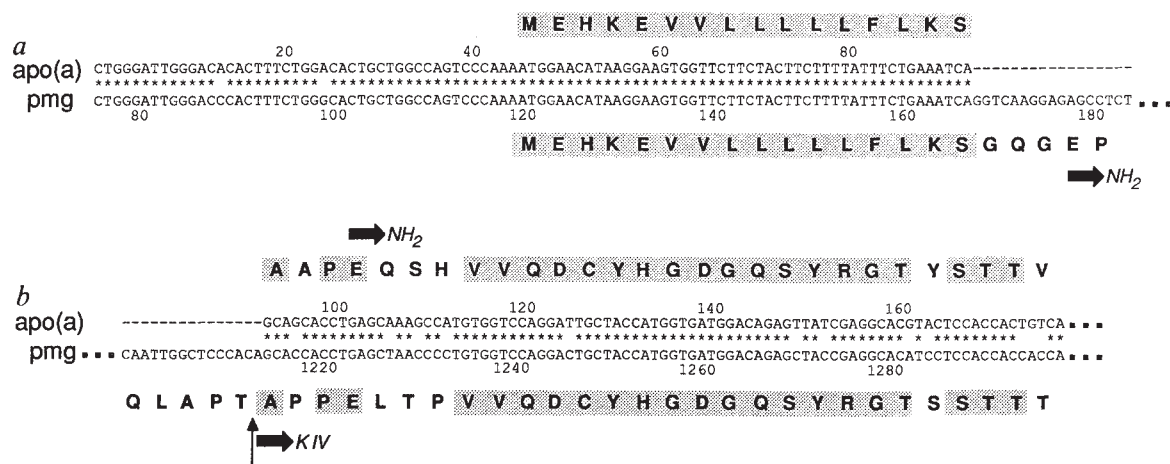
Evolution of the apo(a) gene

The genes for apo(a) and plasminogen are closely related and probably arose from recent gene duplication and subsequent deletion of the exons encoding the amino-terminal 'tail' and kringles 1–3 of plasminogen. The various untranslated and coding sections of the human apo(a) gene range in identity of 78–100% to the human plasminogen gene (Fig. 4). A minimum divergence time for the two genes can be estimated on the basis of a comparison of non-coding sequences in the β -globin regions of human and chimpanzee chromosomes, which suggested that differences accumulate in non-coding sequences at a rate of about 0.3% per million years²³. The homology of the 3' untranslated regions of apo(a) and plasminogen genes suggest that the postulated duplication of a common precursor gene occurred about 40 million years ago, during primate evolution. This estimate, based on the accumulation of unselected DNA changes, refers only to the last time that the two loci exchanged DNA sequence information, either by the initial duplication or by subsequent unequal crossover or gene conversion events. The 5' untranslated regions of the two genes are so nearly identical that a divergence time of less than 7 million years is predicted. This raises the possibility that a gene conversion event involving the 5' sequences occurred after the initial duplication. Inspection of various animal species is needed to address these issues more directly.

Further evidence of the recent evolution of apo(a) comes from the observation that bovine and human plasminogen genes are more divergent than are the human plasminogen and human apo(a) sequences. The partial cDNA sequence of bovine and human plasminogen share 82% homology in the protease domain and 60% in the 3' untranslated region¹⁸; the apo(a) versus human plasminogen homology for these regions is 94% and 87%, respectively. This again implies that the apo(a) gene arose by duplication of (or last encountered recombination with) the plasminogen gene during more recent mammalian evolution.

Although no function of apo(a) is yet known, there are several indications that there has been selective pressure acting upon the gene. Most obvious, of course, is the fact that an extremely long open reading frame of 13,644 nucleotides exists. In the kringle regions, all of the cysteines involved in the three intra-kringle disulphide bonds are conserved, whereas homology is notably diminished in the 'spacer' regions between the highly structured kringle cores (see Fig. 5). Furthermore, the 4-fold repeated kringle differs from the 24-fold repeated kringle in only three nucleotides. Each of these changes is a silent substitution which leaves the encoded amino acids unchanged. In this regard, it is curious that the 'control' regions of this gene may have been adversely affected. Whereas the level of plasminogen is relatively invariant, as are the plasma levels of Lp(a) within an individual, Lp(a) levels differ by ~600-fold among different individuals. This is far greater than the range in concentration of other known components of lipoproteins.

Plasminogen itself is a member of a protein superfamily composed of regulatory proteases of the fibrinolytic and blood coagulation systems²⁴. Duplication, divergence, and exon shuffling has been proposed to account for the presence of



homologous domains in these proteins, which include plasminogen, prothrombin, urokinase, tissue plasminogen activator, coagulation factors XII, VII, IX, X, and protein C (ref. 24). The first five of these proteins contain kringle structures. The protease domains of these proteins all share homology to the archetypal serine protease, trypsin. In general, the function of the non-catalytic segments of these proteins is to mediate the binding to other molecules and hence regulate the cascades of fibrinolysis and blood coagulation²⁵.

The most striking feature of the apo(a) gene is the virtually unprecedented number, length and sequence identity of intragenic repeats. Although other genes contain internally repeated sequences, none yet characterized shows quite this degree of nearly identical repetition²⁶⁻³². The degree of homology of the apo(a) kringle 4-like repeated domains is evidence of very recent duplications (and deletions) caused by out-of-register homologous recombination. Such events are responsible for intragenic repeats as well as the evolution of

Fig. 1 Structure and sequence of human apolipoprotein(a) cDNA clones. *a*, The top line depicts apo(a) mRNA, with the direction of transcription from left to right. The line at the 5' end is untranslated sequence, which is followed by the signal peptide (hatched), and then numbered repeats of plasminogen kringle 4-like coding blocks. Light grey blocks (2-23, 25-26) are identical 342 nucleotide repeats; darker shaded blocks (24, 27-29) are identical to each other but differ by 3 nucleotides from the previous repeats. Kringle 33 contains a deletion of 24 nucleotides and kringle 36 contains a unique cysteine residue. Unshaded, numbered blocks are kringle 4-like domains which differ from the most common kringle ranging from 11 to 74 nucleotides. At the 3' end are the plasminogen kringle 5-like (V) and protease (filled block) coding domains and the 3' untranslated region (line). Below the depicted mRNA are the overlapping cDNA clones that were sequenced. Based on an mRNA size of 14 kb, we propose the existence of 22($\pm$ 2) identical tandem kringle repeats. Owing to the identity of these repeat sequences, the position of clone λ a6 is ambiguous, and the possibility exists that up to three kringle units remain unsequenced. *b*, The amino acids and nucleotides of human apo(a) are numbered at the left of each line. The predicted protein sequence is shown above the DNA. Numbers of the amino acids are negative 1-19 for the predicted signal peptide, and 1 (E)-4529 for the mature protein. OC* denotes the stop codon. The final nucleotide shown is the site of the poly(A) tail, and the presumed polyadenylation signal, AATAAA, is double underlined. Potential asparagine-linked glycosylation sites are marked by asterisks. Amino acid number 4308 (S), corresponds to the arginine in plasminogen which is cleaved to yield proteolytically active plasmin. Precise repeats of the 342 bp DNA sequences are indicated ([A]_n and [B]_n). Amino acids 1-4,207 are homologous to plasminogen kringle 4; residues 4,208-4,308 are homologous to kringle 5; and 4,309-4,529 are homologous to the protease. Also indicated is residue 4,057 (O), a potential free cysteine, and residues 4,350, 4,393 and 4,479 (H, D, S), which correspond to the catalytic triad of plasmin. Nucleotides 80-109 (.....) correspond to the synthetic 30 base oligonucleotide used to isolate 5' clones.

Methods. The first apo(a) cDNA clone, λ a1 was isolated from a randomly-primed cDNA library constructed from liver mRNA isolated from five individuals. All other cDNA clones shown here were subsequently isolated from liver mRNA of a single individual. cDNAs were synthesized using oligo(dT) or random hexanucleotide priming^{41,42}. Size selected cDNAs were ligated into λ gt10 or λ ZAP (clone λ a25; Stratagene) bacteriophage vectors, and screened without amplification⁴³. λ a1 was recovered with a 75-base oligonucleotide probe (PL.1), corresponding to nucleotides 1,306–1,380 of human plasminogen cDNA (ref. 19). The homology of PL.1 to the sequence contained in λ a1 was 82%. Filters were hybridized in $5\times$ SSC (SSC 150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 6.7), $5\times$ Denhardt's solution, 20% formamide, 10% dextran sulphate and $20\ \mu\text{g ml}^{-1}$ of boiled, sonicated salmon sperm DNA at 42°C overnight and washed for 1 h in $2\times$ SSC, 0.1% SDS at 55°C and exposed to X-ray film. All other hybridizations were either under identical conditions or, where indicated, differed only in their concentration of formamide, or the concentration of SSC and temperature of the wash solution. All other clones were obtained by probing libraries with insert fragments of pre-existing clones in 50% formamide and washed in $0.1\times$ SSC at 68°C. Clones λ a11, 4 and 18 were identified as 5' apo(a) clones by rescreening hybridizing plaques with a 30-base oligonucleotide in 25% formamide and washing in $2\times$ SSC at 50°C. This probe was designed to span the coding regions for the end of the apo(a) prepeptide and its mature amino terminus, thus crossing the region of plasminogen that is absent in apo(a). All clones shown here were sequenced to completion except for λ a4, 18 and 9, for which only ~600 bp at both ends were sequenced. DNA sequence analysis was performed by the dideoxy chain termination method⁴⁴ either directly on double-stranded DNA (ref. 45) or single-stranded templates in M13 or pUC118/119 vectors⁴⁶, some of which were shortened by controlled exonuclease digestion⁴⁷.

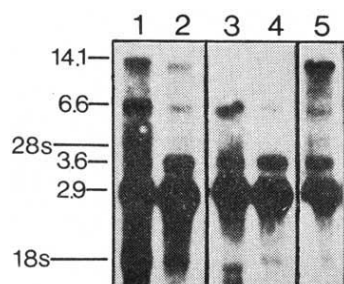


Fig. 3 Northern blot hybridization of apo(a) mRNA. 5 μ g poly(A)⁺ mRNA from human liver (used to construct cDNA libraries) and HepG2 (hepatocarcinoma) cells were electrophoresed in a 1.5% agarose, 6% formaldehyde gel, transferred to nitrocellulose, and hybridized as follows: lanes 1 and 2 contain liver and HepG2 RNA respectively, hybridized with ³²P-labelled RNA probes, clones λ a1 and λ a41, which code for 5' untranslated and signal sequences, several kringle 4-like domains, kringle 5, protease and 3' untranslated sequences. Lanes 3 and 4 are exposures of the same filter shown in lanes 1 and 2, after probe-stripping and decay, and rehybridization with a probe derived from an 866 bp *EcoRI/RsaI* restriction fragment of cloned human plasminogen cDNA which contains the coding sequences of kringles 1, 2 and 3, which are not closely homologous to sequences in apo(a). Note that all but the 14.1 kb band rehybridize. Lane 5 is from the same filter as lane 4, rehybridized without probe-stripping with apoB-100 cDNA to serve as a size marker. The size of hybridizing bands are indicated on the left.

Methods. Poly(A)⁺ mRNA was prepared and electrophoresed as described⁴³. Lanes 1 and 2 were hybridized with an RNA probe derived from the insert of λ a1 plus a 2 kb fragment of λ a41 (kringle 35 to end) subcloned into pSP64 and pSP65 (ref. 48). Lanes 3 and 4 are derived from the same filter after 6 weeks to let ³²P decay and further stripping of probe by washing in 75% formamide, 0.1 $\times$ SSC, 0.1% SDS for 3 h at 42 $^{\circ}$ C. Probe derived from an 866 bp fragment of plasminogen cDNA subcloned into pSP64. Riboprobe hybridizations followed the methods of Fig. 1 with 50% formamide at 60 $^{\circ}$ C and washing in 0.1 $\times$ SSC at 68 $^{\circ}$ C. Lane 5 shows the filter used for lane 4 reprobed without stripping with a labelled 3.5 kb fragment of apoB-100 cDNA, known by sequencing to be 14.1 kb (refs 21, 22). Ribosomal RNA size markers were also used.

multigene families³³. The presence of regions of duplicated sequence retaining ~100% identity implies frequent, if not recent, expansion and contraction of a locus, whose effects can even be seen in populations of the same species. This has been dramatically demonstrated in humans³⁴: among 25 persons studied, individuals contained as few as one, and as many as six, tandemly arranged haptoglobin-related genes. Individuals also have variable numbers of tandemly arranged, nearly identical α -globin genes, γ -globin genes and green eye pigment

genes. Intragenic recombination within repeated sequences is a likely cause of size polymorphism of the fibroin protein, and presumably of other proteins with repetitive amino acid sequences. We therefore propose that the size isoforms of human apo(a) may be due to individual variation in the number of kringle 4-like repeats in the apo(a) gene and protein. Variable carbohydrate content is unlikely to account for the nearly twofold variation in the size of apo(a) isoforms¹⁶.

Suggested structure-function relationships

The homology of apo(a) to plasminogen is distinct from all other apolipoproteins. Seven of the other nine plasma apolipoproteins that have been sequenced are related to each other by the presence of similar amphipathic alpha-helical regions (apolipoproteins A-I, A-II, A-IV, C-I, C-II, C-III, E). Apolipoproteins B and D appear to be unrelated to that group. Although the precise function of all of these proteins is unknown, apolipoproteins B and E serve as receptor ligands, apoC-II is a cofactor for the enzyme lipoprotein lipase, and apoA-I is a cofactor for lecithin-cholesterol acyltransferase. The function of apo(a) is presently unknown, but its homology to plasminogen suggests several functions which can be tested experimentally. The most obvious is that apo(a) could have proteolytic activity. The protease domain of apo(a) contains 88% amino acid identity to plasminogen, with one deletion of 9 residues in apo(a), corresponding to plasminogen residues 716-724. The three amino acids in the catalytic triad are retained. However, the arginine at which plasminogen is cleaved by tissue-type (t-PA) and urokinase-plasminogen activators is substituted by serine in apo(a) as the result of a single nucleotide change (Fig. 1b). To date we have found that neither t-PA, urokinase, or streptokinase can activate apo(a) to obtain active plasmin activity, nor does Lp(a) protein bind ³H-diisopropyl fluorophosphate (which binds the catalytic site of active serine proteases), even after treatment with these three activators of plasminogen. This substitution at the activation site has been seen in two individuals (one from protein and one from cDNA sequencing), and Lp(a) isolated from seven persons cannot be cleaved or activated by t-PA and urokinase. Still, the possibility exists that some individuals could contain a proteolytically active apo(a).

Some of the kringles of plasminogen and other kringle-containing proteins bind to macromolecules during the processes of clot formation and lysis. Plasminogen kringle 1 binds fibrin strongly through a lysine affinity site, whereas kringles 2-5 have reduced or no affinity³⁵⁻³⁸. An immediate possibility was that apo(a) would bind to fibrin by virtue of its many kringle domains. Binding of cholesterol-rich Lp(a) particles to macromolecules exposed at the sites of vessel damage could then contribute to the development of atherosclerosis. We have shown that Lp(a) can bind to lysine-Sepharose¹⁷. Preliminary results, however, show that Lp(a) only binds fibrin weakly, if

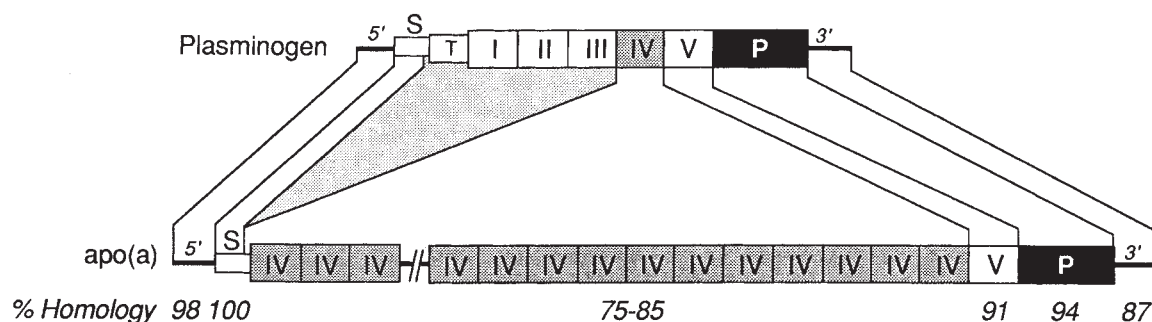


Fig. 4 Sequence structure comparison of plasminogen and apolipoprotein(a). Top line, plasminogen cDNA; bottom line, apo(a) broken into domains. Connecting lines indicate regions of homology with shading to represent the domains lacking in apo(a). The percentage identity of plasminogen and apo(a) cDNA sequences for the following domains is shown at the bottom: 5' untranslated, signal peptide (first 16 codons), kringle 4, kringle 5, protease, 3' untranslated. Other symbols are: T, 'tail' region and plasminogen kringles I, II, III.

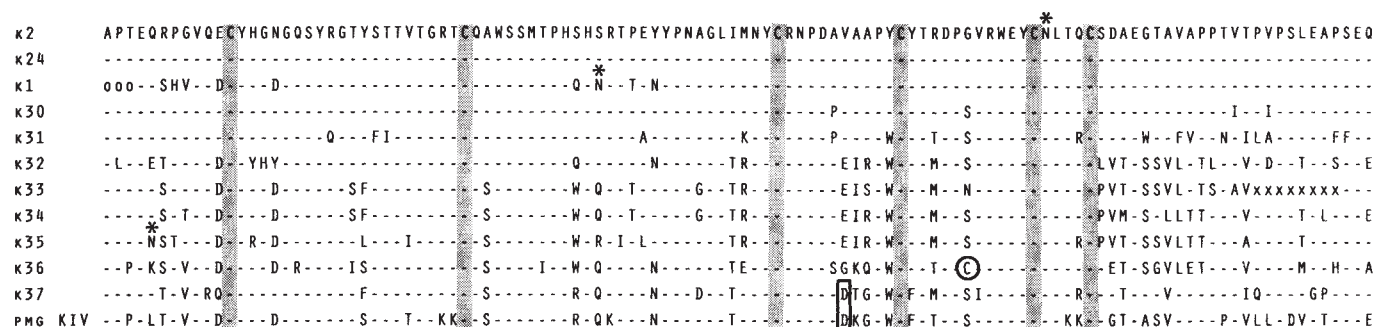


Fig. 5 Kringle 4 sequence comparisons. The amino acid sequences of the kringles of apo(a) and kringle four of human plasminogen (bottom line) are aligned. Dashes denote residues identical to k2, the 24-fold repeated kringle. Kringle 24 is repeated four times in apo(a). The location of cysteine residues involved in intrakringle disulphide bonds in plasminogen are indicated by shading. Also indicated are possible N-linked glycosylation sites (*) present in apo(a), but not in plasminogen kringle 4, and the aspartic acid of plasminogen implicated in fibrin binding and present only in kringle 37 of apo(a) (□). Another feature unique to only one of the apo(a) kringle repeats is a cysteine residue in kringle 36 (○) which might be unpaired in apo(a) and involved in disulphide linkage to apoB-100. X's denote a deletion in kringle 33. Note that most of the sequence divergence occurs in the 'spacer' region which flanks the disulphide linked kringle core. Sequences homologous to the first three residues of kringle 1 (shown by 000) precede the cleavage site of the signal peptide.

at all (D. E., manuscript in preparation). The relatively weak fibrin-binding ability of plasminogen kringle 4 could be further reduced by two relative changes in apo(a). Each of the 37 kringle 4-like repeats of apo(a) contains a possible N-linked glycosylation site, Asn-Leu-Thr, which substitutes for Asn-Leu-Lys at the corresponding location in plasminogen (Fig. 5). There are no N-linked glycosylation sequences in plasminogen kringle 4, and glycosylation at this location in apo(a) might interfere with fibrin binding. In addition, an aspartic acid residue of plasminogen kringle 4 (boxed in Fig. 5) may be directly involved in fibrin binding³⁹. This amino acid is preserved in only one of the 37 apo(a) kringle 4-like repeats. It should also be noted that one of the repeated kringles contains a unique cysteine residue (Fig. 5). As plasminogen contains no free cysteines, and as all other cysteine residues are conserved in apo(a), the cysteine in kringle repeat number 36 may be involved in the covalent linkage of apo(a) to apoB-100.

The most significant effect of apo(a) could be on the receptor-mediated uptake of lipoproteins by one or more of the receptors which bind LDL, modified LDL, or other lipoprotein particles. The covalent linkage of a ~500K apo(a) molecule to apoB-100 would hardly be expected to leave its receptor-binding properties

unaltered. Recently, it has been shown that removal of apo(a) from Lp(a) particles results in a lipoprotein with greatly enhanced affinity for the LDL receptor⁴⁰. At this point, there remain many unanswered questions about Lp(a). The availability of the cloned cDNA has provided suggestive details of the primary structure of apo(a) and leads directly to studies of the apo(a) gene in subjects containing varied plasma levels of Lp(a) and different apo(a) size and sequence polymorphisms. In addition, DNA probes provide additional tools for genetic analysis of Lp(a), the evolution of the apo(a) gene, and the association with cardiovascular disease. But none of the data now available address the issue of a 'useful' function of apo(a). As neither the possible atherogenicity of Lp(a), nor its usefulness as a correlate of disease susceptibility, would provide selective pressure for its evolution, the most significant function of this unusual protein is still to be recognized.

We thank the Genentech DNA synthesis group, Karen Fisher for RNA samples, Gordon Vehar for project support, and Socorro Cuisia and Carol Morita for manuscript preparation. This work was supported by Genentech, Inc., and by NIH. The apo(a) cDNA sequence has been submitted to the Genbank/EMBL databases under accession number M17399.

Received 8 September; accepted 12 October 1987.

- Berg, K. *Acta path. microbiol. scand.* **59**, 369-382 (1963).
- Groener, J. E. M. & Kostner, G. M. *J. lipid Res.* **28**, 1053-1056 (1987).
- Albers, J. J., Adolphson, J. L. & Hazzard, W. R. *J. lipid Res.* **18**, 331-338 (1977).
- Berg, K., Dahlen, G. & Frick, M. H. *J. clin. Genet.* **6**, 230-235 (1974).
- Dahlen, G., Berg, K., Gillnäs, T. & Ericson, C. *J. clin. Genet.* **7**, 334-341 (1975).
- Dahlen, G. *et al. Circulation* **74**, 758-765 (1986).
- Armstrong, V. W. *et al. Atherosclerosis* **62**, 249-257 (1986).
- Költringer, P. & Jürgens, G. *Atherosclerosis* **58**, 187-198 (1985).
- Rhoads, G. G., Dahlen, G., Berg, K., Morton, N. E., Dannenberg, A. L. *JAMA* **256**, 2540-2544 (1986).
- Brown, M. S. & Goldstein, J. L. *Science* **232**, 34-37 (1986).
- Gaubatz, J. W., Heideman, C., Gatto, A. M., Morrisett, J. D. & Dahlen, G. H. *J. biol. Chem.* **254**, 4582-4589 (1983).
- Utermann, G. & Weber, W. *FEBS Lett.* **154**, 357-261 (1983).
- Fless, G. M., Rolih, C. A. & Scanu, A. M. *J. biol. Chem.* **259**, 11470-11478 (1984).
- Gaubatz, J. W., Chari, M. V., Nava, M. L., Guyton, J. R. & Morrisett, J. D. *J. lipid Res.* **28**, 69-79 (1987).
- Fless, G. M., ZumMallen, M. E. & Scanu, A. M. *J. biol. Chem.* **261**, 8712-8718 (1986).
- Utermann, G. *et al. J. clin. Invest.* **80**, 458-465 (1987).
- Eaton, D. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 3224-3228 (1987).
- Malinowski, D. P., Sadler, J. E. & Davie, E. W. *Biochemistry* **23**, 4243-4350 (1984).
- Forsgren, M., Råden, B., Isrealsson, M., Larsson, K. & Heden, L.-O. *FEBS Lett.* **213**, 254-260 (1987).
- Rabiner, S. R., Goldfine, I. D., Hart, A., Summaria, L. & Robbins, K. C. *J. Lab. clin. Med.* **74**, 265-273 (1969).
- Chen, S.-H. *et al. J. biol. Chem.* **261**, 12918-12921 (1986).
- Blackhart, B. D. *et al. J. biol. Chem.* **261**, 15364-15367 (1986).
- Maeda, N., Bliska, J. B. & Smithies, O. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5012-5016 (1980).
- Patthy, L. *Cell* **41**, 657-663 (1985).
- Patthy, L., Trexler, M., Vali, Z., Banyai, L. & Varadi, A. *FEBS Lett.* **171**, 131-136 (1984).
- Benveniste-Schrode, K. *et al. J. mol. Evol.* **22**, 209-219 (1985).
- Wharton, K. A., Johansen, K. M., Xu, T. & Artavanis-Tsakonas, S. *Cell* **43**, 567-581 (1985).
- Koide, T., Foster, D., Yoshitake, S. & Davie, E. W. *Biochemistry* **25**, 2220-2225 (1986).
- Eckert, R. L. & Green, H. *Cell* **46**, 583-589 (1986).
- Gage, L. P. & Manning, R. F. *J. biol. Chem.* **255**, 9444-9450 (1980).
- Manning, R. F. & Gage, L. P. *J. biol. Chem.* **255**, 9451-9457 (1980).
- Boguski, M. S., Birkenmeier, E. H., Elshourbogy, N. A., Taylor, J. M. & Gordon, J. I. *J. biol. Chem.* **261**, 6398-6407 (1986).
- Maeda, N. & Smithies, O. *A. Rev. Genet.* **20**, 81-108 (1986).
- Maeda, N., McEvoy, S. M., Harris, H. F., Huisman, T. H. J. & Smithies, O. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7395-7399 (1986).
- Lerch, P. G. & Rickli, E. E. *Biochim. biophys. Acta* **625**, 374-378 (1980).
- Lerch, P. G., Rickli, E. E., Lergier, W. & Gillissen, D. *Eur. J. Biochem.* **107**, 7-13 (1980).
- Wiman, B. & Collen, D. *Nature* **272**, 549-550 (1978).
- Lucas, M. A., Fretto, L. J. & McKee, P. A. *J. biol. Chem.* **258**, 4249-4256.
- Trexler, M., Vali, Z. & Patthy, L. *J. biol. Chem.* **257**, 7401-7406 (1982).
- Armstrong, V. W., Walli, A. K. & Seidel, D. *J. lipid Res.* **26**, 1314-1323 (1985).
- Okuyama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161-170 (1982).
- Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
- Drayna, D. *et al. Nature* **327**, 632-634 (1987).
- Sanger, F., Nicklen, S. & Coulson, A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Chen, E. Y. & Seeburg, P. H. *DNA* **4**, 165-170 (1985).
- Vieria, J. & Messing, J. *Meth. Enzym.* (in press).
- Dale, R. M. K., McClure, B. A. & Houchins, J. P. *Plasmid* **13**, 31-40 (1985).
- Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035-7056 (1984).

Excess infrared radiation from a white dwarf—an orbiting brown dwarf?

B. Zuckerman* & E. E. Becklin†

* Department of Astronomy, University of California, Los Angeles, California 90024, USA

† Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

We have discovered that the white dwarf star Giclas 29–38 appears to emit substantial radiation at wavelengths between 2 and 5 μm , far in excess of that expected from an extrapolation of the visual and near-infrared spectrum of the star. The infrared colour temperature of the excess radiation is $1,200 \pm 200$ K and, at the distance of G29–38, corresponds to a total luminosity of 5×10^{-5} solar luminosities ($L_{\odot}$). If the excess 3.5- μm radiation is emitted by a single spherical body at 1,200 K, then its radius is 0.15 solar radii ($R_{\odot}$). These characteristics are similar to those that have been calculated for substellar objects called brown dwarfs. The most natural interpretation of our observations is that there is a substellar, somewhat Jupiter-like brown dwarf in orbit around G29–38.

The history of searches for planets in orbit around stars other than the Sun has been long and rather chequered. Various search techniques, both optical and infrared, have been employed. Perhaps the most famous is that of classical optical astrometry where one looks for a periodic wiggle in the path of a star across the sky. The wiggle of the visible star is in response to the gravitational attraction of an unseen planet. There have been various claims in the scientific and popular literature of discovery of planets or planet-like bodies in orbit around other stars, but none has ever been confirmed¹. In addition, there do exist a few cool isolated objects whose discoveries have been confirmed and which may be substellar^{1,2}. (A substellar object is one that does not fuse protons into helium nuclei in its interior.) But because these isolated objects are not in orbit around stars, they cannot, in any event, be called planet-like.

There is thus at the moment not a single confirmed example of an extrasolar object, either isolated or in orbit around a star, that is unambiguously substellar.

We are involved in an observational programme to search for cool companions to white dwarf stars by means of infrared photometry. Because white dwarfs are faint at near-infrared wavelengths, even a moderately cool companion ($T \geq 1,000$ K) can be detected in reasonable integration times. The first targets in our survey were the white dwarfs in the Hyades and Pleiades, which are young, nearby star clusters. A discussion of the negative results of that search, and other background material, has been published in ref. 3. We are now surveying field white dwarfs, and results of that search will be reported subsequently. The purpose of this letter is to describe our data on the most interesting of the field white dwarfs examined to date, Giclas 29–38 (also called WD 2326+049, ZZ Psc, Gl 895.2, EG 159, and LTT 16907). We use the Giclas, Burnham and Thomas nomenclature, because their finding charts in the Lowell Observatory Series have been invaluable in survey work of this sort. The physical properties of G29–38 are listed in Table 1; it has been classified as spectral type DAV4 in the scheme of Sion *et al.*⁴.

All the data shown in Fig. 1 and Table 1 were obtained under clear skies in the early morning hours (local time) on Monday 24 August 1987 at the Mauna Kea Observatory in Hawaii, with the 3-m NASA Infrared Telescope Facility (IRTF). In fact, the infrared excess reported here was first detected in the early morning on 23 August but thin patchy clouds were present and the measurement accuracy was therefore somewhat degraded. Nonetheless, the data from the two days agreed to within $\sim 10\%$. We employed the standard IRTF InSb photometric system, with

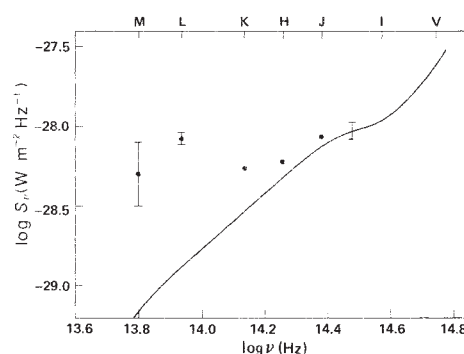


Fig. 1 Plot of flux density against frequency for G29–38. The black dots are the measured fluxes. The line is the flux that a star with the effective temperature of G29–38 might be expected to have based on the average fluxes measured for six of our programme stars other than G29–38. The vertical bar on the line between J and I is a typical standard deviation in the colour about the mean of the six stars. The vertical bars on the M and L points are one standard error of the mean. The errors in the K, H, and J fluxes are equal to the size of the black dots. The width of the infrared bandpasses are of order 0.4 μm , except for L which is ~ 1 μm wide.

apertures of various angular diameters in the telescope focal plane. The secondary mirror was chopped by 20 arc s at 7 Hz between the white dwarf and a reference position, while the telescope was nodded every 10 s so that the white dwarf appeared first in one beam and then in the other.

Figure 1 displays the total flux received from the direction of the white dwarf at wavelengths between 5,500 Å and 4.8 μm . The solid line is our best estimate of the flux that would be received at each wavelength from the white dwarf alone. We estimated this expected flux by using six of our programme stars with temperatures within a few thousand degrees of the temperature of G29–38. It is obvious from Fig. 1 that there is huge excess emission at all wavelengths ≥ 2 μm . Based on our observations and those of other astronomers, we know that extrapolation of the visual spectra of white dwarfs into the infrared is very predictable, and reliably follows the Rayleigh–Jeans tail of a hot black body. The only exceptions are a few white dwarfs that have close, low mass, M-type main sequence companions^{3,5}. The influence of the M stars on the combined white dwarf plus M star spectrum is already evident at 8,000 Å (ref. 6) and the M star totally dominates the combined spectrum by 1.25 μm wavelength. By contrast, there is no hint (see Fig. 2 in ref. 6) of anything amiss in the 8,000 Å brightness of G29–38, and even at 1.65 μm the total flux barely deviates from the flux expected from a white dwarf at 11,500 K (Fig. 1). The excess infrared emission from G29–38 is therefore significantly cooler than that due to any known M-type main sequence star.

We believe that the excess infrared emission is very unlikely to be due to a background object that happens to lie near the line of sight to G29–38. We mapped the distribution of 1.25 and 3.5 μm emission toward G29–38 with an ~ 2 arc s diameter aperture. At both wavelengths the majority of the emission originates from a single localized region and the position of the maximum emission at the two wavelengths agrees to better than 1 arc s. As the 1.25- μm emission is due to the white dwarf (Fig. 1) the source of the excess 3.5- μm emission is, therefore, coincident with the white dwarf to 1 arc s. The probability that a random, highly red-shifted background galaxy or a distant, intrinsically rare, dust enshrouded giant star should coincide so precisely with the white dwarf must be minuscule. In addition, we examined the Palomar Sky Survey (PSS) plates in the vicinity of G29–38. Because of its fairly large proper motion, the position at which G29–38 now lies is clearly visible on the PSS plates, uncontaminated by light from G29–38. Down to the plate limit, there is no indication of any star or galaxy at the

Table 1 Measured and derived quantities for G29–38

Band	V	I	J	H	K	L	M	References
Wavelength (μm)	0.55	0.80	1.25	1.65	2.2	3.5	4.8	
Apparent total magnitude*	12.95	13.43	13.03 ± 0.05	13.03 ± 0.05	12.65 ± 0.05	11.31 ± 0.10	11.2 ± 0.4	6, IRTF
White dwarf (apparent magnitude)†	12.95	13.43	13.24 ± 0.15	13.28 ± 0.15	13.35 ± 0.15	13.35 ± 0.15	13.35 ± 0.15	6, see Fig. 1
Brown dwarf (apparent magnitude)				14.7 ± 0.5	13.45 ± 0.2	11.5 ± 0.1	11.4 ± 0.5	see Fig. 1
Brown dwarf (flux density/mJy)				1.2 ± 0.7	2.6 ± 0.6	7.0 ± 0.8	4.4 ± 2.0	
1950 position of G29–38 (Epoch 1987.7)				$23 \text{ h } 26 \text{ min } 15.0 \text{ s} \pm 0.1 \text{ s } 4^{\circ}58'25.6'' \pm 1.5''$				IRTF
Distance to G29–38					$14.1 \pm 0.7 \text{ pc}$			18
Temperature of G29–38					11,500 K			9
Luminosity of G29–38					$\sim 2 \times 10^{-3} L_{\odot}$			
Temperature of brown dwarf					$1200 \pm 200 \text{ K}$			see text
Luminosity of brown dwarf					$5 \times 10^{-5} L_{\odot}$			see text
Radius of brown dwarf					$0.15 R_{\odot}$			see text
Mass of brown dwarf					$0.04\text{--}0.08 M_{\odot}$			see text

* Assumes Vega = 0 mag at all wavelengths.

† Errors from variation among magnitudes of six programme stars with similar effective temperatures.

current position of G29–38.

If the excess infrared radiation is associated physically with G29–38, then it is probably due to either a swarm of tiny dust particles heated by G29–38, or a single, roughly Jupiter-sized brown dwarf. Dust particles in the vicinity of stars are found either in an outflowing cool wind, in an accretion flow, or in an orbiting disk. Although some hot white dwarfs (with $T_{\text{eff}} \geq 30,000 \text{ K}$) are associated with outflowing gas, driven presumably by radiation pressure in resonance lines, we know of no example of any radiation pressure mechanism that would work effectively to drive matter from a white dwarf with $T_{\text{eff}} = 11,500 \text{ K}$. G29–38 is a pulsating ZZ Ceti star^{7,8} and pulsations appear to be an effective contributor to rapid mass loss from Mira-type variable red giant stars. However, the ZZ Ceti phenomenon is believed⁹ to be due to high order g-mode pulsations that appear mainly as variations in the effective stellar temperature at essentially constant radius. Such pulsations will not lead to mass loss.

A more plausible way to explain the infrared excess is to invoke the existence of an orbiting disk of particulate matter around G29–38. In particular, we envisage a white dwarf analogue of the Vega or Beta Pictoris main sequence phenomenon discovered¹⁰ by the IRAS satellite. The effective temperature of the excess infrared radiation ($\sim 1,200 \text{ K}$) implies that the hypothesized radiating dust particles must be located at a distance $\sim 1 R_{\odot}$ from G29–38. This temperature is about the maximum that can be tolerated by silicate dust grains before they evaporate. G29–38 is located far from any known interstellar dust clouds. The depletion time due to the Poynting–Robertson effect $1 R_{\odot}$ from a star of $2 \times 10^{-3} L_{\odot}$ for particles with radii of $0.5 \mu\text{m}$ is only $\sim 10 \text{ yr}$ (ref. 11), which is very short compared to the time necessary to travel from an interstellar cloud to the present position of G29–38. Therefore, the source of dust grains must be intimately associated with G29–38.

What might be a plausible source of dust grains this close to the remnant of a star that underwent a red giant expansion not quite 10^5 yr ago? Occasional cometary impacts onto white dwarf stars may explain certain photospheric spectroscopic peculiarities¹² and it has been suggested recently that near-misses of comets and white dwarfs could effectively produce circumstellar gas in orbit around the white dwarf (M. Jura, F. Coroniti and C. Alcock, in preparation). Possible evidence for the existence of such circumstellar gas is described in ref. 13 for four white dwarfs that are all substantially hotter than G29–38. The minimum mass of dust grains required to explain our observations is $\sim 3 \times 10^{17} \text{ g}$, which could be supplied by six typical solar system comets^{12,14}. Accumulating this much orbiting material in such close proximity to G29–38 seems very

unlikely given the rapid depletion due to the Poynting–Robertson effect and the absence of any spectral peculiarities in its photospheric spectrum (J. Greenstein, personal communication) which might be expected¹² as a consequence of the rapid accretion of the orbiting material.

A more attractive possibility is that a warm brown dwarf is in orbit around G29–38. The upper mass limit for brown dwarfs in general, is reasonably precisely defined by the lower mass end ($0.08 M_{\odot}$) of the proton burning main sequence¹⁵. The lower mass limit, which represents also the upper mass limit to large planets, is much less precisely defined. We prefer $0.01 M_{\odot}$ because this is the lowest mass object that can burn deuterium into helium¹⁵ and also because much below this mass, support against gravitational collapse is supplied in significant part by the electromagnetic force in addition to the electron degeneracy pressure that supports brown dwarfs—and white dwarfs such as G29–38 as well.

In Table 1 we summarize the observed and derived properties of the G29–38 system if it is composed of a white dwarf and a brown dwarf. The time for G29–38 to cool to 11,500 K is $\sim 6 \times 10^8 \text{ yr}$ (refs 16, 17). The black body that best fits our observations has a temperature of 1,200 K. Using this black body, the measured $3.5\text{-}\mu\text{m}$ flux, and the known distance (14.1 pc) to G29–38¹⁸ we derive the luminosity and radius given in Table 1. If the putative brown dwarf companion was unaffected by the previous red giant phase of the evolution of G29–38 and if the progenitor of G29–38 had a main sequence mass near the maximum permitted for stars that become white dwarfs ($\sim 6 M_{\odot}$), then the minimum cooling age for the brown dwarf is $\sim 6 \times 10^8 \text{ yr}$, and its minimum mass is $0.04 M_{\odot}$ according to Figs 2 and 3 in Nelson *et al.*¹⁵. Their Fig. 1 implies a corresponding brown dwarf radius of $\sim 0.1 R_{\odot}$, which is $\sim 50\%$ smaller than we deduce (Table 1) for $T = 1,200 \text{ K}$. Lower mass main sequence progenitors of G29–38 would imply slightly larger masses for the brown dwarf and slightly smaller radii, thus increasing slightly the apparent discrepancy between theory and observation.

The discrepancy between our derived radius and the theoretical value might be explained away by one or more of the following effects: errors in the photometry and calibration ($\sim 15\%$); errors in the distance to G29–38 ($\sim 5\%$); or underestimation of the effective temperature of the brown dwarf due to the presence of strong H_2O bands or other gaseous or particulate matter in its atmosphere¹⁹. In particular, if the actual temperature of the brown dwarf is 1,500 K, then the derived radius is in reasonable agreement with Fig. 1 in ref. 15. In addition, there might be errors in the theoretical models.

We note that models by Livio and Soker²⁰ imply that low-mass brown dwarfs will usually be transformed into low-mass stars if they are engulfed in the atmosphere of their red giant primary during the asymptotic giant branch of stellar evolution. It seems to us conceivable that such a mechanism might occasionally transform a low-mass brown dwarf into a higher mass brown dwarf. In this case, the cooling clock for the brown dwarf companion to G29–38 could have started at the same time that G29–38 itself became a white dwarf ($\sim 6 \times 10^8$ yr ago).

Finally, we consider the possibility that the infrared excess could be due to a true planet, such as Jupiter, that revolves at a distance of $\sim 0.5 R_\odot$ from G29–38 and is heated to $\sim 1,200$ K by absorption of starlight. The pulsational spectrum of G29–38 is very complicated (ref. 8 and R. E. Nather, personal communication) and the exact low-frequency periodicities that would be produced by a companion star are difficult to extract from the data. Based on a very limited data set, McGraw and Robinson⁸ did suggest a possible companion with an orbital period of just under two hours which would be consistent with a planet (or brown dwarf) at the appropriate distance from G29–38. However, the rotational period of such a planet would almost certainly be synchronous with its orbital period and the planet would have a hot and a cold hemisphere. This would appear in our observations as an obvious secular variation in the 2.2- μ m magnitude. No variation has been observed, implying a planetary orbit nearly in the plane of the sky if this model is correct. An additional, more severe, problem is that a planet located $0.5 R_\odot$ from G29–38 that initially had the size of Jupiter would overfill its Roche lobe and would transfer a substantial fraction of its mass onto G29–38.

The most natural explanation for the excess infrared radiation in the direction of the white dwarf G29–38 appears to be a substellar brown dwarf in orbit around G29–38. An apparently less likely, but perhaps not impossible, explanation of the observation is a disk of orbiting particulate matter around G29–38. Only future observations with higher spatial and spectral resolution than those reported here can decide between these possibilities.

We thank Dr Jesse Greenstein for valuable discussions, Dr David Morrison for critical support to this experiment. We thank Patricia Wheeler for typing the manuscript and other important assistance, and Drs Michael Jura, Donald Hall, Len Cowie, and Harland Epps for helpful comments. This research was supported in part by grants from the NSF to UCLA and the University of Hawaii. B.Z. and E.E.B. are visiting astronomers at the Infrared Telescope Facility, operated by the University of Hawaii under contract to NASA.

Note added in proof: K. Hodapp and M. Rigler have obtained 1.65 and 2.2 μ m images of G29–38 using a 64×64 element IR/CCD with 0.14 arc s pixels on the University of Hawaii 2.2 m telescope. The 0.8 arc s image (FWHM) appears point-like, thus implying a maximum projected separation of ≤ 5 AU between the white dwarf star and any companion.

Received 11 September; accepted 9 October 1987.

A giant intergalactic H I bubble near Arp143

P. N. Appleton*, F. D. Ghigo†, J. H. van Gorkom‡, James M. Schombert§ & Curtis Struck-Marcell||

* School of Physics and Astronomy, Lancashire Polytechnic, Preston PR1 2TQ, UK

† Department of Astronomy, University of Minnesota, Minnesota 55455, USA

‡ National Radio Astronomy Observatory, VLA Project, PO Box 0, Socorro, New Mexico 87801, USA

§ Palomar Observatory, California Institute of Technology, Pasadena, California 91125, USA

|| Astronomy Program, 12 Physics Building, Iowa State University, Ames, Iowa 50011, USA

We present observations of the nearby ring galaxy Arp143 (NGC2445/4) and describe the discovery of a giant H I shell 20 kpc across, lying 100 kpc from Arp143. The shell, which shows evidence for line splitting near its centre, lies in a long H I filament pointing towards the Arp143 system. We suggest that the shell is an expanding bubble of H I, and is direct evidence for a powerful galactic-scale explosion 10^8 years ago which we associate with a strong burst of star formation. We believe that both the ring galaxy and the giant bubble have their origin in the passage of a gas-rich companion through the inner disk of NGC2445, and its subsequent disruption.

The H I observations were made during December 1986 and March 1987 with the C and D arrays of the Very Large Array (VLA); full details of the data reduction will be given elsewhere. Figure 1 shows a contour map of the integrated H I emission superimposed on a Johnson V-band image of Arp143 and its surroundings. A neutral hydrogen plume $150(100 \text{ km s}^{-1} \text{ Mpc}^{-1}/H_0)$ kpc in length (assuming a distance for Arp143 of 40 Mpc; H_0 is Hubble's constant) extends from Arp143 northwards. The H I distribution is almost unresolved with our present beam near the base of the plume, but appears to flare out about two-thirds of the way along it. The flaring is caused by the existence of an incomplete shell-like structure embedded in the plume and centred at RA(1950) = 7 h 43 min 23 s and Dec(1950) = $39^\circ 18.0'$. Its H I mass is $\sim 3 \times 10^8 M_\odot$. Its shell-like nature is most striking in Fig. 2, a radio-photo of the emission, centred near the velocity of the bubble. This partly incomplete shell (diameter = 20 kpc) is almost circular and is brightest to the west, where it appears to be embedded in the side of the H I plume. The shell is not empty but contains weak filamentary H I. Overall it resembles a limb-brightened shell similar to the H I shells discovered around old supernova remnants¹, but on a vastly larger scale. An enlargement of the optical image in the shell region shows a small galaxy located on the edge of the shell (see Fig. 1b) with a Holmberg diameter of 12.5 arc s (2.4 kpc if at the distance of Arp143) and an apparent magnitude of $m_v = 19.69$. It appears to have the colours of a blue dwarf irregular galaxy ($B - V = 0.46$ mag). The galaxy may be associated with the shell.

The bubble-like nature of the shell is further reflected in the velocity structure of the shell region. Figure 3 shows a selection of H I spectra across the shell. The dominant emission comes from the H I plume which runs from the south-east to the north-west, but within and near the centre of the shell weak double and in some cases multiple broad profiles are seen. The brightest component is red-shifted with respect to the velocity component of the plume by $\Delta V = 180 \text{ km s}^{-1}$, and is seen over approximately a beam area (28×28 arc s) near the geometrical centre of the shell structure. These results suggest that the shell is a bubble of gas expanding and compressing the surrounding H I material. The expansion in the radial direction seems to be asymmetric, because one of the velocity components is similar

- Liebert, J. & Probst, R. G. *Astr. Astrophys.* **25**, 473–519 (1987).
- Probst, R. G. & Liebert, J. *Astrophys. J.* **274**, 245–251 (1983).
- Zuckerman, B. & Becklin, E. E. *Astrophys. J.* **319**, L99–L102 (1987).
- Sion, E. M. *et al. Astrophys. J.* **269**, 253–257 (1983).
- Probst, R. G. *Astrophys. J. Suppl.* **53**, 335–349 (1983).
- Greenstein, J. L. *Astrophys. J.* **276**, 602–620 (1984).
- Shulov, O. S. & Kopatskaya, E. N. *Astrofizika (SSR)* **10**, 117–119 (1973).
- McGraw, J. T. & Robinson, E. L. *Astrophys. J.* **200**, L89–L93 (1975).
- Holm, A. V. *et al. Astrophys. J.* **289**, 774–781 (1985).
- Aumann, H. H. *et al. Astrophys. J.* **278**, L23–L27 (1984).
- Wyatt, S. P. & Whipple, F. L. *Astrophys. J.* **111**, 134–141 (1950).
- Alcock, C., Fristrom, C. C. & Siegelman, R. *Astrophys. J.* **302**, 462–476 (1986).
- Bruhweiler, F. C. & Kondo, Y. *Astrophys. J.* **269**, 657–667 (1983).
- Delsemme, A. H. in *Comets* (ed. Wilkening, L. L.) 85–230 (University of Arizona, Tucson, 1982).
- Nelson, L. A., Rappaport, S. A. & Joss, P. C. *Astrophys. J.* **311**, 226–240 (1986).
- Iben, I. & Tutukov, A. V. *Astrophys. J.* **282**, 615–630 (1984).
- Mazzitelli, I. & D'Antona, F. *Astrophys. J.* **308**, 706–720 (1986).
- Harrington, R. S. & Dahn, C. C. *Astr. J.* **85**, 454–465 (1980).
- Berriman, G. & Reid, N. *Mon. Not. R. Astr. Soc.* **227**, 315–329 (1987).
- Livio, M. & Soker, N. *Mon. Not. R. Astr. Soc.* **208**, 763–781 (1984).

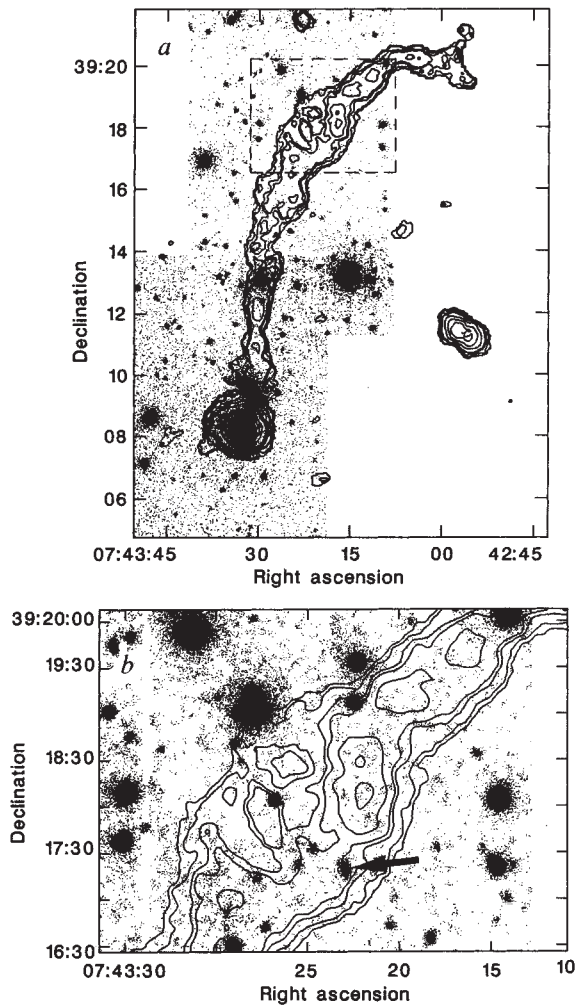


Fig. 1. *a*, Integrated H I emission from the Arp 143 region superimposed on two V-band images taken with the Palomar 1.5-m RCA CCD system⁴. The VLA data was obtained with both the C and D arrays. After continuum subtraction, the channels containing lines were CLEANed⁵ and combined resulting in a map with a resolution of 28×28 arc s. Contours are 18, 40, 76, 100, 181, 272, 363, 726, $1,090 \text{ Jy ms}^{-1}$ per beam. *b*, The boxed region of *a* enlarged, showing more details of the optical image in the bubble region. The arrow shows a faint blue galaxy projected against the western edge of the shell. Most of the other objects in the field are believed to be foreground stars.

to the H I velocity in the adjacent plume. The present data are consistent with an expanding bubble having burst from the side of the H I plume, accelerating clouds to high velocities. We estimate that $3 \times 10^7 M_{\odot}$ of H I resides in this high-velocity component. This implies that a mechanical energy of $M\Delta V^2/2 = 9.7 \times 10^{47} \text{ J}$ would be required to give rise to the observed high velocity gas alone.

The nearly circular shape of the bubble sets a purely kinematic upper limit to its age. The H I plume in which the bubble is embedded shows a regular velocity gradient along its length of $\sim 1.5 \text{ km s}^{-1} \text{ kpc}^{-1}$, which can be inverted to give a characteristic local plume lifetime of $6.6 \times 10^8 \text{ yr}$. If the bubble had grown to its present size in a time longer than this it would be stretched in the direction of the plume. This is not the case, so the bubble must be younger than $7 \times 10^8 \text{ yr}$, a timescale consistent with an expansion age of the bubble of $1.1 \times 10^8 \text{ yr}$ determined by dividing the diameter of the bubble by the velocity of the observed line-splitting.

We believe that the shell structure is a relic of a very large



Fig. 2 A grey-scale 'radio-photo' of the sum of three velocity channels centred at $V_{\odot} = 3,935 \text{ km s}^{-1}$, illustrating the filamentary shell-like nature of the bubble region.

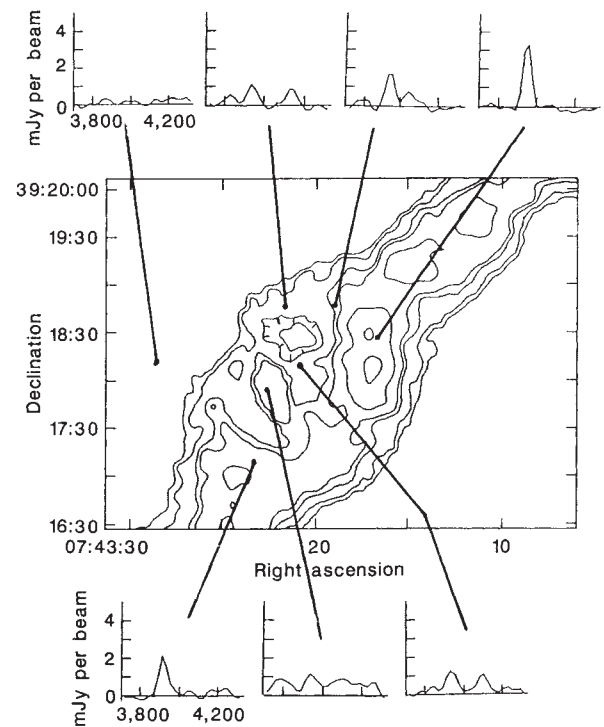


Fig. 3 A selection of H I spectra of the bubble region showing the weak but broad profiles inside the H I shell. The strongest feature is redshifted relative to the main emission from the plume by 180 km s^{-1} .

explosion. A conservative estimate of the energy needed to form the bubble is simply the mechanical energy seen in the redshifted H I near the bubble centre (see above), which would correspond to the energy of 10^4 supernova explosions of $\sim 10^{44} \text{ J}$. This is almost certainly an underestimate by a factor of 10, because the total shell mass is approximately ten times that of the redshifted component.

We will discuss in detail elsewhere the origin of the H I plume and simply list here three possible explanations. The plume may be (1) gas stripped by ram pressure in a hot intergalactic corona

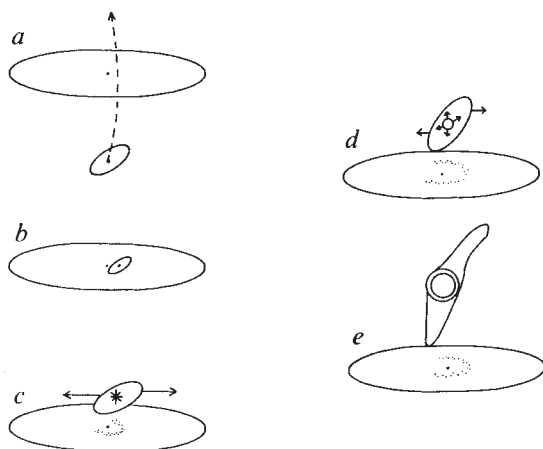


Fig. 4 A possible explanation for the origin of the H I bubble involving the near-central collision between an H I-rich galaxy and NGC2445. In a schematic time sequence of events (a-e), a starburst is triggered in the intruder (c), which grows to form a shell. A ring-like density wave forms in the disk of NGC2445 (see text).

from the companion NGC2444, (2) a tidal tail formed by a close encounter of NGC2444 with NGC2445, or (3) debris from a companion galaxy which has plunged through the disk of NGC2445 causing the formation of a ring-shaped density wave within the disk. Only (3), outlined below, provides a link between the formation of the H I plumes, the existence of the giant H I bubble and the ring-like nature of NGC2445.

Models suggest² that a central collision between an intruder galaxy and a disk galaxy would lead to ring formation in the disk. The most obvious contender for such an intruder is NGC2444, the early type ($B-V=1.03$) galaxy lying along the minor axis of NGC2445. The first difficulty with this picture is that there is no H I emission directly associated with NGC2444 ($V_0 = 3,997 \pm 50 \text{ km s}^{-1}$; J. Huchra, personal communication), despite its similar systematic velocity to NGC2445 ($V = 4,010 \pm 50 \text{ km s}^{-1}$). It could be argued that NGC2444 was originally a gas-rich galaxy, but has recently plunged through the disk of NGC2445 and been stripped of gas. But then the stripped gas would trail behind the galaxy, and the fact that the plume only begins to the north of NGC2444 and extends away from NGC2445 makes this idea implausible.

A mechanism which is consistent with the collisional ring galaxy picture involves the disruption of an H I-rich galaxy of low surface brightness during its passage through the central regions of NGC2445. This is shown schematically in Fig. 4. The dwarf system is disrupted by dynamical heating³ as it passes through the disk plane of NGC2445, and the debris of both the stars and the gas is spread over a wide range of velocities and positions approximately one free-fall time later (Fig. 4c, d). We assume here that the gas in the companion has sufficient momentum to punch a hole in the gas disk NGC2445. The stellar component of the intruder galaxy is then spread over a considerable area on the sky and would be difficult to detect. This is consistent with the detection of a very faint optical component to the plume in the south (not shown here). If, at the time of impact (Fig. 4b), a starburst is triggered near the nucleus of the intruder, then an expanding shell would later form by the action of supernovae and stellar winds (Fig. 4c). As the intruder galaxy moved out of the potential well of NGC2445 and was dynamically disrupted, the shell would expand to its present size in a low density environment. Apart from triggering a starburst in the intruder galaxy, the slightly off-centre collision between the galaxies would drive a ring-like density wave into the disk of NGC2445 (Fig. 4c-e) which is consistent with the appearance and kinematics of the gas and stars in this galaxy (data to be published elsewhere).

Whatever the origin of the H I plume, the discovery of the H I bubble provides strong evidence for the sudden release of mechanical energy in interacting galaxies. In this case the shell of gas may provide a site for the subsequent formation of dwarf galaxies from the ejecta.

The VLA of the National Radio Astronomy Observatories is operated by Associated Universities Inc., under contract with the US NSF.

Received 31 July; accepted 21 September 1987.

1. Braun, R. thesis, Univ. Leiden (1985).
2. Toomre, A. in *IAU Symp. No. 79*, 109-116 (1978).
3. Tremaine, S. in *The Structure and Evolution of Normal Galaxies* (ed. Fall, S. M. & Lynden-Bell, D.) 67-84 (Cambridge University Press, 1981).
4. Schombert, J. M. & Wallin, J. F. *Astrophys. J.* (in the press).
5. Hoggom, J. A. *Astr. Astrophys. Supp.* 15, 417-426 (1974).

The neutrino emission of SN1987A

R. Schaeffer, Y. Declais* & S. Jullian†

Service de Physique Théorique, CEN-Saclay,
91191 Gif-sur-Yvette Cedex, France

* Centre de Physique des Particules de Marseille,
Faculté des Sciences de Luminy, 13288 Marseille Cedex 9, France
† Laboratoire de l'Accélérateur Linéaire, Centre d'Orsay,
Bt 200, F 91405 Orsay, France

The appearance of a supernova in the Large Magellanic Cloud made it possible for the first time to detect the neutrino emission from the explosion of a stellar object. Several groups reported neutrino events, a few hours before the onset of the optical signal¹⁻⁴. The Mont Blanc collaboration¹ has reported a signal that the Kamioka² detector, supposedly more sensitive, has not seen. The Kamioka², IMB³ and (marginally) the Baksan⁴ groups have reported a positive detection that occurred five hours later. Here we show that the assumption that the Mont Blanc (or Baksan) observation is due to the formation of a neutron star leads to excessive energy requirements. If confirmed, these two detections would invalidate any of the presently proposed explanations for the SN1987A event. The Kamioka and IMB observations, on the other hand, match perfectly the standard theory and we suggest that they were the tracers of the neutron star formation. They can then be used to constrain the number of neutrino flavours to be less than four.

The standard mechanism for a star explosion leaving a neutron star is the collapse of the iron core of a massive star⁵⁻¹⁰. During the initial phase of the collapse, a sizeable fraction of the protons transform into neutrons⁹, with the emission of a ν_e of energy $\leq 10 \text{ MeV}$. The collapse phase lasts until the infalling matter becomes opaque to neutrinos. A few $10^{56} \nu_e$ are emitted during this phase. As the collapsed core reaches nuclear matter densities, an outgoing shock develops. When the shock reaches the less dense layers that are still transparent to neutrinos, another $\sim 10^{56} \nu$, mostly ν_e , are expected to be emitted. The second of these bursts produces neutrinos with an average energy $\sim 10 \text{ MeV}$ ⁵⁻¹⁰. The whole process lasts much less than one second. The remainder of the gravitational energy ($\sim 2 \times 10^{53} \text{ ergs}$) is emitted subsequently in the form of $\nu\bar{\nu}$ pairs of all flavours (e, μ, τ). Although the neutron star contains $\nu\bar{\nu}$ pairs of very high energy (100 MeV), only the low energy ones are emitted because the neutrino mean free path is strongly energy dependent. The energy is larger for the μ and τ neutrinos for which only neutral currents contribute whence their larger average emission energy. In most estimates⁶⁻¹⁰ the $\nu_e\bar{\nu}_e$ pairs are of about 10 (possibly 15) MeV and the energy of the other species is around 20 MeV. The theoretical estimate of these neutrino energies relies on the calculation of their mean free path in nuclear matter. Although their weak interaction with

Table 1 Binding energies and surface redshifts of model neutron stars

Equation of state	$M = 1.4 M_{\odot}$		$M_{\max}$		
	$E_B(\text{erg})/(1+z_s)$	z_s	$E_B(\text{erg})/(1+z_s)$	$M(M_{\odot})$	z_s
PN	2.3×10^{53}	0.30	3.2×10^{53}	1.7	0.54
BJI	2.0×10^{53}	0.23	3.5×10^{53}	1.9	0.50
(MFT)	1.9×10^{53}	0.18	6.7×10^{53}	2.8	0.61
(PN+a)	3.0×10^{53}	0.64	3.0×10^{53}	1.4	0.64
Observed		0.2			0.2

The binding energy E_B and surface redshift z_s were obtained¹³ from realistic (PN²⁰, BJI²¹) equations of state, or from those incorporating extreme parameters (MFT²², PN+a¹⁶) that do not reproduce all classical properties of nuclei and dense matter. The results of ref. 16 for the 'standard' neutron star have been rescaled to $M = 1.4 M_{\odot}$ which implies a $\sim 20\%$ correction, using the techniques developed in ref. 19. Note that the observation of the surface redshift is a severe constraint on the available binding energy.

electrons and nucleons is well understood, this is a difficult many-body problem. Measured energies that are different from those expected would not necessarily imply that a dramatic revision of the standard neutron star models is needed. The duration of the emission, also, is very uncertain, with estimates ranging from 1 to 100 s, but is typically taken to be about 10 s. This expectation leaves little room for the observation of two bursts separated by a five hour delay, as seems to be implied by measurements of the Mont Blanc experiment on the one hand, and those of the Kamiokande-IMB-Baksan experiments on the other. While this problem remains unresolved, we thus proceed by assuming that there was only one burst.

To study the implications of the neutrino observations, an estimate of the neutrino flux that they imply is necessary. The expected flux can be easily obtained using energy conservation. Indeed, for a given neutron star binding energy E_B and an average neutrino energy $\bar{E}_\nu$, and distance D to the Large Magellanic Cloud, this flux is

$$F \approx 4.2 \times 10^{10} \left(\frac{E_B}{2 \times 10^{53} \text{ erg}} \right) \left(\frac{\bar{E}_\nu}{10 \text{ MeV}} \right)^{-1} \left(\frac{D}{50 \text{ kpc}} \right)^{-2} \text{ cm}^{-2} \quad (1)$$

for all neutrinos. Assuming equipartition of energy between all $\nu\bar{\nu}$ pairs, the flux for one species (say ν_e or $\bar{\nu}_e$) is reduced sixfold. Equipartition is expected to be a fair approximation for neutrinos whose mass is much less than 10 MeV. Because of the occurrence of the two ν_e bursts before the $\nu\bar{\nu}$ emission, and also because of the higher average energy of the μ and τ neutrinos, the ν_e or $\bar{\nu}_e$ flux is somewhat smaller than this estimate, and also somewhat model dependent. Early calculations showed that the reduction factor may be as large as 10, but the more recent estimates⁶⁻⁸ all show it to be rather close to six. Using the latter value gives a flux for ν_e or $\bar{\nu}_e$

$$F_e \approx 7.0 \times 10^9 \left(\frac{E_B}{2 \times 10^{53} \text{ erg}} \right) \left(\frac{\bar{E}_\nu}{10 \text{ MeV}} \right)^{-1} \left(\frac{D}{50 \text{ kpc}} \right)^{-2} \text{ cm}^{-2} \quad (2)$$

Assuming D and $\bar{E}_\nu$ can be directly measured, the binding energy E_B can be obtained from the number of detected neutrinos. Let us emphasize that this binding energy is rather well determined by the current models for neutron stars. The measured neutrino flux will thus provide very stringent constraints on these models. Indeed this binding energy, which has to be calculated using general relativity, is basically the gravitational energy GM^2/R of the star, and is thus free of most of the uncertainties inherent in estimates relying on strong interactions. Moreover, the gravitational mass M of neutron stars ranges in the rather narrow 1 to $2 M_{\odot}$ interval. This energy (see table 1) is $E_B \approx 2 \times 10^{53}$ erg for a typical $1.4 M_{\odot}$ neutron star for all realistic equations of state, extreme values-for cases already excluded by experiments on nuclei-ranging from 1.6 to 3×10^{53} . For neutron stars with the maximum possible mass, E_B may reach 3.5×10^{53} erg but this requires the unrealistic assumption

Table 2 Comparison between observed and expected neutrino events

	Reported		Expected for $E_\nu = 10 \text{ MeV}$	
	$N_\nu + N_{\bar{\nu}}$	$\langle E_\nu \rangle$	$\mathcal{M}$ (tonnes)	$N_\nu + N_{\bar{\nu}}$
Mont Blanc	5	7.6	90	0.4
Baksan	3	9.2	130	0.6
Kamiokande	11	15	2,140	8
IMB	9	32	5,000	18*

The comparisons assume that a neutron star with standard ($M = 1.4 M_{\odot}$) properties has been formed. The average energy of the observed neutrinos is $\langle E_\nu \rangle$, and $\mathcal{M}$ is the effective mass of the detector. The expected counts are given for a 100% detector efficiency. The efficiency correction is discussed in the text.

* This number is to be substantially reduced because of the high energy threshold imposed in this experiment.

that the neutron star is formed happens by chance to have just this mass. Moreover, the surface redshift z_s (that is roughly $z_s \sim GM/Rc^2$, although it has a more complex expression in general relativity) has been measured in one case, $z_s \approx 0.2$. If this measurement were possible for SN1987A, it would represent a very stringent constraint on the possible mass and binding energy of the corresponding neutron star.

In the energy range under consideration here, the charged current interactions are only energetically possible for $\bar{\nu}_e$ neutrinos, via the $\bar{\nu}_e p \rightarrow e^+ n$ reaction. Due to the large mass of the proton, the positron escapes with nearly all the energy of the incoming $\bar{\nu}_e$ ($E_{e^+} = E_{\bar{\nu}_e} - 1.8 \text{ MeV}$), but the memory of the direction of the latter is lost. The neutrino elastic scattering on electrons $\nu(\bar{\nu})e \rightarrow \nu(\bar{\nu})e$ proceeds via the neutral as well as the charged current interaction. In contrast with the scattering on protons, the escaping electron is emitted in the direction of the neutrino, but with much lower energy $E_{e^-} \approx E_\nu/2$, on average.

To calculate the number of $\bar{\nu}_e p$ interactions, only the scattering on hydrogen protons is important. The $\bar{\nu}_e$ interaction with the ^{12}C and ^{16}O atoms in the target, whose nucleons are tightly bound, is negligible for $E_\nu < 30 \text{ MeV}$ (^{12}C) and 45 MeV (^{16}O). The standard value^{12,13} of the $\bar{\nu}_e p$ cross-section can, in the kinematical range considered, be approximated by $\sigma_{\bar{\nu}_e p} \approx 7.5 \times 10^{-44} E_\nu^2 (\text{MeV})$, to obtain a first estimate of the number of expected $\bar{\nu}_e p$ events for a given neutron star binding energy E_B , neutrino average energy $\bar{E}_\nu$, distance D , detector mass $\mathcal{M}$:

$$N_{\bar{\nu}_e p} = K \left(\frac{E_0}{2 \times 10^{53} \text{ erg}} \right) \left(\frac{\bar{E}_\nu}{10 \text{ MeV}} \right) \left(\frac{D}{50 \text{ kpc}} \right)^{-2} \left(\frac{\mathcal{M}}{1000 t} \right) \epsilon \quad (3)$$

with $K = 4.5$ for a H_{2n}C_n scintillator, and $K = 3.5$ for water. The ratio $\epsilon = E_0^2/\bar{E}_\nu^2$ depends on the model used for the neutrino spectrum. Typical values are, however, very close to 1 and range from $\epsilon \approx 0.8$ for realistic neutrino energy distributions to 1.3 for a black-body thermal distribution¹⁴. A sizeable reduction of equation (3) is to be expected when the detector efficiency to be considered in detail below is taken into account.

The ν_{e^-} and $\bar{\nu}_{e^-}$ scattering events are due to both neutral and charged current interactions. The sum of the ν and $\bar{\nu}$ events of all flavours leads to a number of expected counts

$$\sum_{\nu\bar{\nu}} N_{\nu\bar{\nu}} = 0.46 \left(\frac{E_B}{2 \times 10^{53} \text{ erg}} \right) \left(\frac{D}{50 \text{ kpc}} \right)^{-2} \left(\frac{\mathcal{M}}{1000 t} \right) \quad (4)$$

if we use the Weinberg-Salam model to estimate the cross-sections¹⁵, $\sigma_{\nu_e e} \approx 0.96 E_\nu (\text{MeV}) \times 10^{-44} \text{ cm}^2$, $\sigma_{\bar{\nu}_e e} \approx 0.41 E_\nu (\text{MeV}) \times 10^{-44} \text{ cm}^2$ and $\sigma_{\nu_{\mu, \tau} e} \approx \sigma_{\bar{\nu}_{\mu, \tau} e} \approx \sigma_{\nu_e e} \approx \sigma_{\bar{\nu}_e e}/7$. These estimates assume that there is no reduction in the counts due to events that are missed because of threshold effects or various other detector limitations. They lead to the ratio

$$N_{\bar{\nu}_e p} / \left(\sum_{\nu\bar{\nu}} N_{\nu\bar{\nu}} \right) = 9.7 (\bar{E}_\nu / 10 \text{ MeV}) \quad (5)$$

The latter gets much higher if the detector efficiency is taken

into account: detailed estimates using realistic energy distribution of the neutrinos show in all cases that the $\bar{\nu}e^-$ counts are much more reduced than the $\bar{\nu}p$ count by a finite energy threshold. The neutrino scattering on electrons thus cannot be detected by the IMB experiments, in contrast to the Kamiokande, Mont-Blanc and Baksan detection for which, however, it requires at most a 10% correction to the counts. This means that most of the directional information is lost in the detection.

We now proceed to consider the absolute counts. The number of expected events is given in Table 2, assuming maximum detector efficiency. As the detector efficiency is necessarily less than 100% the comparison with the reported counts leads to an inequality valid whatever the detector efficiency is:

$$\left(\frac{E_B}{2 \times 10^{53} \text{ erg}}\right) \left(\frac{\bar{E}_\nu}{10 \text{ MeV}}\right) \geq 13 \quad (6)$$

for the Mont Blanc experiment. This is a dramatic constraint, especially since the neutrino average energy is measured and is slightly below 10 MeV. No presently available nuclear model could provide the binding energy $\geq 3 \times 10^{54}$ implied by this measurement of the neutrino flux. There is a similar constraint from the Baksan experiment which is, however, somewhat less severe. The Kamiokande experiment, on the other hand, implies for an effective mass of 2,140 tonnes

$$\frac{E_B}{2 \times 10^{53} \text{ erg}} \frac{\bar{E}_\nu}{10 \text{ MeV}} \geq 1.4 \quad (7)$$

and is thus consistent with the current models for neutron stars. The constraint from the IMB neutrino detection is more difficult to assess, due to the larger detection efficiency correction because of the higher energy thresholds imposed in this experiment. In this case, also, the importance of the $\bar{\nu}e$ interactions with ^{16}O should be investigated.

We now proceed by taking the detector efficiency^{2,3} of the various experiments into account. To extrapolate from the detected flux to the neutrino energy flux that reached the Earth, a model for the neutrino energy distribution is necessary. The simplest possibility would be to take a black-body Fermi-Dirac distribution¹⁴, but realistic models⁷⁻⁸ show some deviations from this first guess, the neutrino mean free path being energy dependent. The estimate of the emitted energy flux is thus somewhat model dependent. This is a negligible effect for the Kamioka observation, but affects the extraction of information from the IMB counts since the latter correspond to the tail of neutrino energy distribution. From the evaluation of the detector efficiency^{2,3} we find that the 11 neutrinos detected by Kamioka imply a total energy extracted by the $\bar{\nu}e$

$$E_{\bar{\nu}e}^{\text{tot}} = 6.5_{-3}^{+2.5} \times 10^{52} \text{ erg} \quad (\text{Kamioka}) \quad (8)$$

and, assuming equipartition of energy and three flavours:

$$E_B^{\text{exp}} = 3.9_{-1.8}^{+1.5} \times 10^{53} \text{ erg} \quad (\text{Kamioka}) \quad (9)$$

These 1σ errors include the uncertainty in the measured neutrino energies and the statistical errors. The corresponding values for the joint analysis of the Kamioka and IMB experiments using the theoretical spectra of Nadëzhin and Ostroshenko⁵ or those of Woosley and Wilson and Mayle⁷ ($15 M_\odot$) that give consistent results, are

$$E_{\bar{\nu}e}^{\text{tot}} = 8.0_{-2.5}^{+2.3} \times 10^{52} \text{ erg} \quad (\text{Kamioka} + \text{IMB}) \quad (10)$$

$$E_B^{\text{exp}} = 4.8_{-1.5}^{+1.4} \times 10^{53} \text{ erg} \quad (\text{Kamioka} + \text{IMB}) \quad (11)$$

These numbers are reduced by 20% if only the eight neutrinos seen by Kamioka during the first two seconds are considered. The IMB data alone would imply an energy $E_{\bar{\nu}e}$ that is much more model dependent, as already emphasized, but consistent with equation (8) at the 1σ level.

From the theoretical estimate of the neutron star binding energy E_B^{th} , still assuming energy equipartition, we derive the

number of flavours

$$N_f = (0.63_{-0.14}^{+0.28}) \frac{E_B^{\text{th}}}{10^{53} \text{ erg}} \quad (\text{Kamioka} + \text{IMB}) \quad (12)$$

For all realistic models of neutron star with standard parameters ($1.4 M_\odot$), E_B^{th} is smaller than 3.5×10^{53} ergs. This leads to a 90% probability that

$$N_f < 3.5 \quad (\text{Kamioka} + \text{IMB}) \quad (13)$$

The latter bound is equal to 4 at the 97% confidence limit only. It obviously scales as $(D/50 \text{ kpc})^{-2}$. The uncertainty of D is difficult to assess, because there are systematic differences in the values quoted by various authors. If we use the high value $D = 52.2 \text{ kpc}$ ²³ we get $N_f < 3.2$ at the 90% confidence level (CL). But for values as low as $D = 46 \pm 3 \text{ kpc}$ ²⁴, $N_f < 4.2$ at the 90% CL and $N_f < 4$ at the 87% CL.

The standard model for stellar collapse leading to a neutron star thus predicts just the energy emission that was observed by the Kamioka and IMB detectors, and it is therefore likely that those detectors indeed detected the neutrino emission of SN1987A. The Mont-Blanc event then remains unexplained. The Kamioka group maintain that, given their detector efficiency, they would have seen a significant signal but did not; thus there is the possibility that the Mont-Blanc event is not due to SN1987A.

A model has been proposed^{25,26} that incorporates the Mont-Blanc observation, suggesting that the Mont-Blanc detection was due to the formation of a metastable neutron star, and the subsequent signal, five hours later, to the formation of a black hole. Our calculation shows that this model has severe difficulties due to the energy requirements of the Mont-Blanc detection if attributed to SN1987A. A metastable neutron star cannot have¹⁷⁻¹⁹ a mass well above the Oppenheimer-Volkov mass, and thus cannot meet the energy requirements. Also, the typical metastability times are expected to be a fraction of a second or even much less¹⁷⁻¹⁹. So, there is no natural explanation for the five hours delay in this case, and this model must invoke a new, *ad hoc*, but very efficient mechanism to explain the time delay between the two bursts. The model is not supported by the SN1987A early light curve^{11,27} whose timing is consistent with the occurrence of the collapse at the time when the second neutrino burst was reported. A final resolution would obviously be provided by the observation of the pulsar remnant implied by the standard model, though not by the black-hole formation model.

We would like to thank D. N. Schramm for discussion.

Received 2 April; accepted 30 September 1987.

1. Aglietta, M. *et al. Europhys. Lett.* **3**, 1315-1320; 1321-1324 (1987).
2. Hirata, K. *et al. Phys. Rev. Lett.* **58**, 1490-1493 (1987).
3. Bionata, R. M. *et al. Phys. Rev. Lett.* **58**, 1494-1496 (1987).
4. Pomansky, A. Moriond meeting, March 5-14, 1987 (Montmerle, T. & Tran Thanh Van, J.) (Frontières Gif-sur-Yvette, France, 1987).
5. Nadëzhin, D. K. & Ostroshenko, I. V. *Astr. Zh.* **57**, 78-89 (1980).
6. Burrows, A. & Mazurek, T. J. *Nature* **301**, 315-318 (1983).
7. Woosley, S. E., Wilson, J. R. & Mayle, R. *Astrophys. J.* **302**, 19-34 (1986).
8. Mayle, R., Wilson, J. R. & Schramm, D. N. *Astrophys. J.* (in the press).
9. Beethe, H. A., Brown, G. E., Applegate, J. & Lattimer, J. M. *Nucl. Phys. A* **324**, 487-533 (1979).
10. Brown, G. E., Beethe, H. A. & Baym, G. *Nucl. Phys. A* **375**, 481-532 (1981).
11. Schaeffer, R., Cassé, M., Mochkovitch, R. & Cahen, S. *Astr. Astrophys.* (in the press).
12. Yamaguchi, K. *Prog. theor. Phys.* **23**, 1117-1132 (1960).
13. Reines, F. & Cowen, C. L. *Jr Phys. Rev.* **113**, 273-279 (1959).
14. Bahcall, J. N., Dhar, A. & Piran, T. *Nature* **326**, 135-136 (1987).
15. Kaiser, B., Fischbach, E., Rosen, S. P. & Spivak, H. *Phys. Rev. D* **20**, 87-98 (1979).
16. Haensel, P. & Proszynski, M. *Astrophys. J.* **258**, 306-320 (1982).
17. Schaeffer, R., Haensel, P. & Zdunik, L. *Nucl. Phys. A* **381**, 519-543 (1982).
18. Haensel, P., Zdunik, L. & Schaeffer, R. *Astr. Astrophys.* **160**, 251-258 (1986).
19. Zdunik, L., Haensel, P. & Schaeffer, R. *Astr. Astrophys.* **172**, 95-110 (1987).
20. Pandharipande, V. R. *Nucl. Phys. A* **178**, 123-144 (1971).
21. Malone, R. C., Johnson, M. B. & Beethe, H. A. *Astrophys. J.* **199**, 741-748 (1975).
22. Haensel, P., Kutschera, M. & Proszynski, M. *Astr. Astrophys.* **102**, 299-302 (1981).
23. Sandage, A. R. & Tamman, G. *Astrophys. J.* **167**, 293-310 (1971).
24. De Vaucouleurs, G. *Astrophys. J.* **223**, 730-739 (1978).
25. Hillebrandt, W. T. *et al. Astr. Astrophys. Lett.* (in the press).
26. De Rujula, A. *Phys. Lett. B* **193**, 514-524 (1987).
27. Arnett, D. *Astrophys. J.* (in the press).

Rapid changes and near-stationarity of the geomagnetic field during a polarity reversal

Carlo Laj, Sylvie Guitton & Catherine Kissel

Centre des Faibles Radioactivités, Laboratoire Mixte CNRS-CEA, Avenue de la Terrasse, BP No. 1, 91190 Gif-sur-Yvette, France

One of the most important suggestions from many of the new reports of sedimentary geomagnetic records is that the time dependence of the changes of the geomagnetic field during a reversal might be more complex than early records indicated. Indeed, alternating periods of rapid and slow field changes are apparent in some sedimentary records¹⁻⁴. In volcanic records, very fast directional changes recorded by individual lava flows during their progressive cooling have been reported^{5,6}, but a major problem, when inferring the temporal behaviour of the transitional field, is that the extrusion rate of the successive lava flows is much more variable and more difficult to evaluate than the rate of accumulation of the sedimentary sequences. Here we report the results of a detailed study of a geomagnetic reversal recorded in Middle Miocene marine clays, with temporal resolution almost comparable with that of the volcanic records and, in our opinion, a more reliable chronological control. The record displays large fluctuations in the rate of directional change and different episodes of collapse and recovery of the field intensity. A very fast N-S directional change may be entirely controlled by a frozen-flux dominated reversal mechanism.

The sampled section, of Serravallian-Tortonian boundary age as determined by nannofossil stratigraphy⁷, is near the village of Maherado on the flank of a major monocline on the island of Zakynthos in Greece. This island has undergone a 26° clockwise rotation and a 5-6° northward drift since the deposition of the sediments^{8,9}. The section consists of a well exposed fresh outcrop, with a very regular bedding plane tilting 30° towards the northeast. Previous magnetostratigraphic work has revealed three polarity reversals at levels 8.8, 11.8 and 17 m from the bottom respectively. Here we report only the results from the lowest one, called ZM01.

The cores were obtained with an electric drill powered by a portable generator. We normally used a 25-mm diameter corer 26 cm long, with the drill installed on a sliding frame perpendicular to the bedding. Cores up to 24.5 cm long were recovered. Multiple measurements of the stratigraphic position of the cores gave similar results (within 0.5-1 cm) between cores, depending on their position. The stratigraphic position of the different samples from each core was calculated taking into account the stratigraphic position of the top of the core, its orientation parameters and even the thickness of the saw used to cut the cores. Each of the very long cores drilled perpendicular to the bedding covered an appreciable fraction of the transitional zone with no relative stratigraphic errors among the different samples from it. In some cases up to 8 specimens 22 mm long, in others 16 specimens 11 mm long ones, or 29 slices 5 mm thick were obtained from a single long core.

X-ray diffraction analyses showed that the mineralogy was essentially detrital quartz, clay minerals, some feldspars and a carbonate fraction. Chlorite, illite and chaolinite dominated the clay fraction. A microprobe examination of a series of samples showed that the concentration of the major elements Ca, Si, Al and Fe, in order of decreasing abundance, was constant over the entire section.

Stepwise acquisition of isothermal remanent magnetization in fields up to 1.8 T indicated that 95±5% of the saturation isothermal magnetization (SIRM) was acquired after 0.2 T. The median destructive field of the SIRM was 25 mT and the maximum blocking temperatures ranged between 500 and

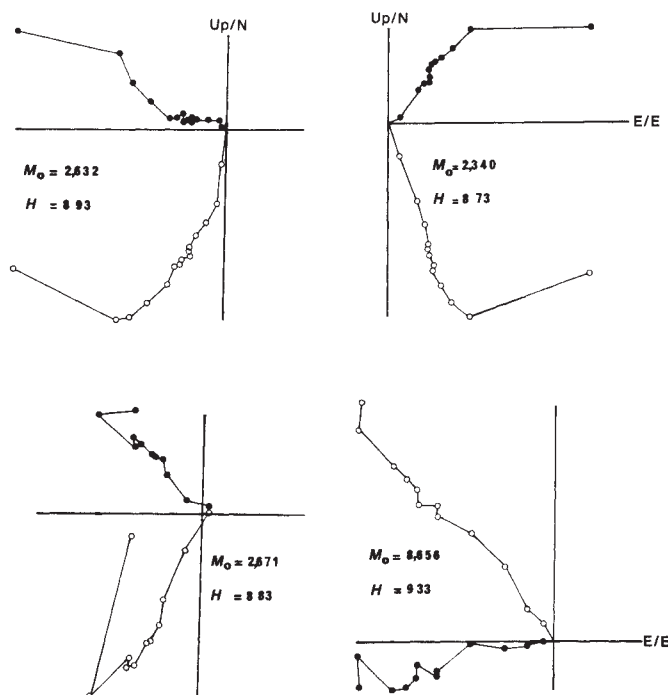


Fig. 1 Typical demagnetization diagrams. The stratigraphic height of the samples, reported in each diagram, identify their position in the stratigraphic columns of Figs 2 and 4. The solid symbols refer to declination data, the open symbols to the projected inclinations. In all the diagrams the fifth demagnetization step corresponds to 300 °C. The NRM intensity, M_0 , is given in units 10^{-6} A m^{-1} .

580 °C. All these results are consistent with a magnetic mineralogy dominated by magnetite. Plots of the anhysteretic susceptibility, χ (anhysteretic remanent magnetization, ARM) against the low-field susceptibility yield a single well-defined regression line, consistent with a uniform grain size of less than 10 μm (refs 10, 11).

The constancy of all the sedimentary and magnetic parameters over the entire section is an indication of a uniform, undisturbed sedimentary environment, so that one may assume that no drastic changes of the sedimentation rate occurred during the deposition of the section. The hypothesis of a constant sedimentation rate is a good first-order approximation.

All the measurements were made with a LETI 3-axes cryogenic magnetometer. Thermal demagnetization gave consistent results and was used throughout. Each sample was demagnetized with a minimum of ten steps from room temperature to the limit of reproducible results. No significant changes of the magnetic susceptibility were observed up to the highest temperatures, indicating that the magnetic minerals were not seriously affected by the thermal treatment. The stable palaeomagnetic directions, isolated after 250-300 °C, were determined from demagnetization diagrams (Fig. 1).

Figure 2 shows the final directions as a function of the stratigraphic height after correction for the tectonic tilt. In these plots each point corresponds to one sample, either a 22 mm × 25 mm or an 11 mm × 25 mm cylinder, except between 8.50 and 8.70 m where 5 mm × 25 mm samples were used, after a thorough check that this shape did not alter the magnetometer reading. The declinations have been corrected by subtracting 26° to account for the rotation of the island. The mean inclination value well outside the transition zone is 45°, ~12° shallower than that expected at this latitude. This arises from a compaction-induced inclination error of ~6-7° already observed in similar sediments from the Hellenic Arc⁸, and from the northward displacement of ~5-6° in latitude⁹. Intriguingly, higher (~50°) inclinations are observed at the end of the transition.

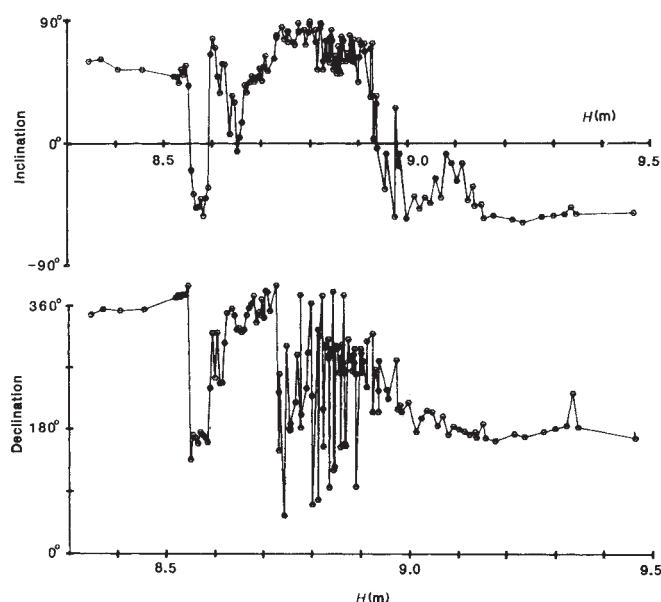


Fig. 2 Plots of the observed inclination and declination values as a function of stratigraphic height. The transition is characterized by a very fast onset. The fluctuations present on the declination record between 8.7 and 8.9 m correspond to a period of very high inclinations. Their amplitude is thus amplified by the projection.

The sedimentation rate, needed for a precise description of the time dependence of the transitional field, can be obtained only if a correlation of the results to the polarity timescale can be established. In spite of the precise biostratigraphic dating at the Serravallian-Tortonian boundary (12 Myr), this correlation is not entirely straightforward, owing to the short time span of the sampled section. Nevertheless, in the interval 12 ± 3 Myr, a single well-defined maximum of the correlation coefficient is observed, using either the Lowrie and Alvarez¹² or the Berggren *et al.*¹³ polarity scales. In both cases this corresponds to correlating the observed magnetostratigraphic column with anomaly 5A, and yields an average sedimentation rate between 2 and 3 cm kyr⁻¹, a very reasonable figure for such sediments in the heliennic sedimentary arc.

The complicated directional pattern of the ZM01 at Maherado is best described in terms of virtual geomagnetic pole (VGP) paths, calculated taking into account the rotation of the island and assuming a palaeoposition 6° S of its present one. A correction of the inclination error with a formula such as that reported by King¹⁴ does not significantly change the results¹⁵, so no correction has been applied.

Figure 3a shows that the reversal starts with a full N-S episode, so fast that intermediate directions are recorded in each core drilled in this zone by only one or two 5-mm samples: the episode thus covers a maximum stratigraphic interval of 10–15 mm. The intermediate directions suggest an ill-defined path east of the site's longitude. A period of standstill at the south pole is then recorded by a few samples, spanning a stratigraphic height of 5–6 cm, after which the VGP moves back to the North pole. The corresponding S-N path is defined by a few samples (total duration 6–7 cm) and is situated approximately on a longitudinal band 90° W of the site.

The VGPs then cluster about the north pole for some time (about 10 cm) and start moving again towards the mid-latitudes more or less along the site meridian (Fig. 3b). In contrast with the rapidity of the first N-S-N cycle, the VGPs then linger at northern mid-latitudes for a relatively long time (30% of the total stratigraphic height of the reversal) with quite a large scatter of the field directions both longitudinally and latitudinally. After this period of standstill the VGPs start moving again first to the west, then southward along a rather narrow longitudinal band

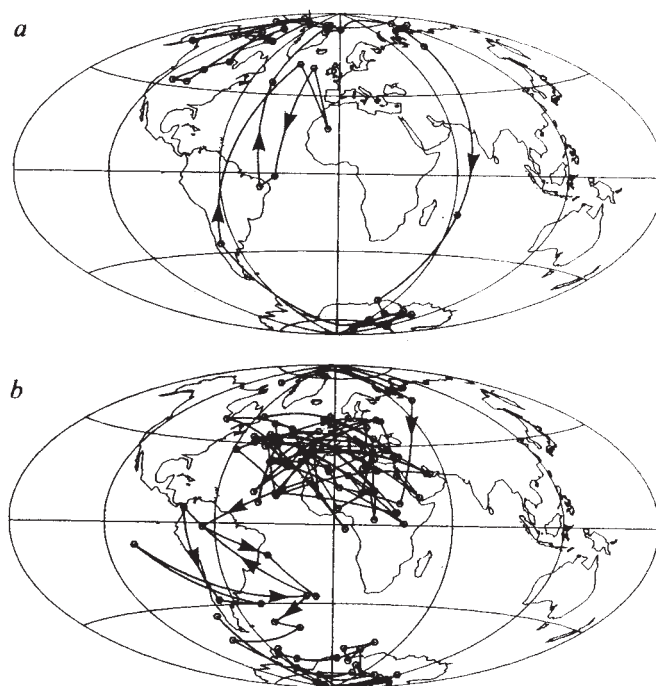


Fig. 3 Plot of the VGP path during the reversal. We have separated the path of first N-S-N cycle (a) from that of the final N-S change (b). This separation is arbitrary and only intended to increase the clarity of the plot. There are very few points on the first N-S-N cycle in spite of the fact that 5-mm samples from the same core were used, indicating a very fast rate of change of the geomagnetic field vector.

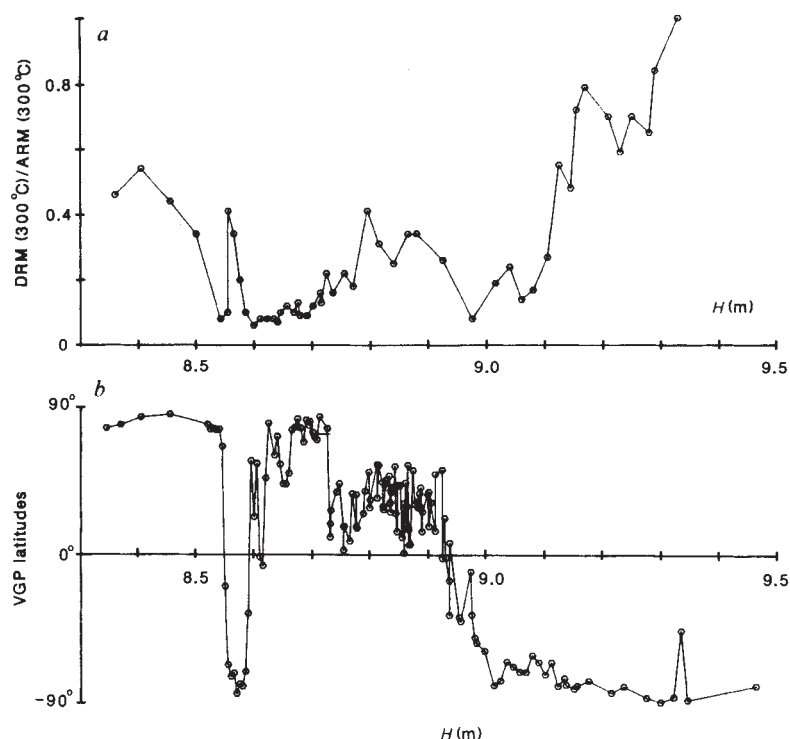
centred on 90° W of the site. The south pole is reached after a final oscillation that brings the VGP back to equatorial latitudes. The total stratigraphic height involved in the directional changes is 50 cm.

The relative changes of the field intensity have been obtained using ARM to normalize the variations of the concentration of the magnetic minerals carrying the remanence^{16,17}. This method has been applied on 'twin' samples, as described elsewhere¹⁵. Figure 4 shows the variations of the ratio detrital remanent magnetization (DRM) (300 °C)/ARM (300 °C), which represents our best approximation to the relative changes of the field intensity, against stratigraphic height, together with a plot of the VGP latitudes.

The initial palaeointensity has about half its post-transitional value, indicating that our sampling does not extend low enough in stratigraphic height and, more fundamentally, that the dipole strength decreased before the onset of the directional changes. There is a sharp drop in field intensity slightly before the beginning of the very fast N-S directional change. Then the intensity is re-established to about half its post-transitional value for the brief period during which reverse dipolar directions are observed. The intensity then drops again while the VGP returns to the north pole and does not recover when the VGP is in a normal (north) dipole field position. The time during which intermediate directions occur (VGPs lingering at mid-latitudes) is also marked by a progressive increase of the field intensity again to about half its post-transitional value. Finally, the field intensity drops again during the last part of the reversal, and is ultimately re-established to non-transitional values after completion of the reversal.

It has been suggested that some of the characteristics of sedimentary records of reversal could result partly from smoothing of the signal because of the complex mechanism of acquisition of remanence¹⁸. However, some of the fine-scale features present in the record show that the influence of a possible post-depositional remanent magnetization (PDRM) is limited.

Fig. 4 Plots of the variation of the relative field intensity (*a*) and of the VGP latitude (*b*) as a function of stratigraphic height. The intensity plot was obtained with 'twin' samples, as described in ref. 15. In the lower part of the transition we have used a core, broken when drilled, for which the orientation parameters were lost but a reliable stratigraphic control was available. For this reason a small discrepancy exists between the VGP latitude data and the intensity data. A significant increase in field intensity occurs after the first N-S directional change, suggesting a temporary re-establishment of the usual (reverse) dipole field. No recovery occurs when the VGP reaches again a stable northern polarity. The partial re-establishment of the field intensity for mid-latitude VGPs indicates a rather stable configuration with a non-dipolar field direction. A few full-polarity VGP latitudes at the end of the transition are intriguingly about 5° higher than those measured well outside the transition zone.



The peak of the intensity recovery after the first N-S episode, for instance, is only 4 cm wide at mid-height, ruling out the possibility of significant smoothing. The presence of the fast N-S episode recorded over a very limited stratigraphic height is also evidence that the record has not been smoothed as a result of a thick lock-in-zone of the magnetization, which would cause blocking of different magnetic grains from the same stratigraphic horizon at different times. We thus believe that the Maherado clays correctly record fast changing field directions, and can be used for the study of geomagnetic reversals.

The N-S segment at the onset of the reversal is ill-defined, so that no information about the geometry of the reversal can be derived from it. The following S-N segment, although not perfectly defined, lies more or less 90° W of the site, indicating that large non-zonal terms are present at this stage in the harmonic content of the transitional field. During the following and final N-S directional change the VGPs lie initially more or less along the site's meridian then, after the period of near standstill, move to a longitude 90° W of the site for the second half of this transition. Thus large changes in the structure of the transitional field occur during this N-S segment, zonal terms probably dominating the onset, whereas non-zonal ones are apparent during the second half.

From a dynamical point of view, the Maherado record is characterized by periods of rapid or very rapid changes alternating with periods of near-stationarity. The first N-S directional change is recorded over a maximum stratigraphic interval of <1.5 or 2 cm, which corresponds to a duration of only 500–1,000 years at the inferred sedimentation rate. This is much shorter than intervals usually reported in the literature but is not unique. A similar duration of about 800 years has recently been reported for a reverse-normal field reversal of Matuyama age recorded in lake sediments in Kashmir¹⁹. Additional evidence that extremely rapid directional changes of the geomagnetic field may occur during a reversal is given by the Steens Mountain record^{5,6}. Thus the very rapid N-S change at Maherado is not inconsistent with some of the known characteristics of the reversing geomagnetic field.

The recovery and successive decay of the field intensity occurring at the end of the first N-S phase when the field direction is due south is also a rapidly evolving phenomenon. Indeed,

the entire process, which we interpret as a temporary re-establishment of a usual (reverse) dipolar configuration, only lasts about 2,000 years. A similar phenomenon has already been observed during the period between the two phases of the Steens Mountain reversal^{5,6}. On the contrary, no recovery of the field intensity is observed during the subsequent period, lasting about 3,000–4,000 years, characterized by normal (northerly) directions. Thus, in spite of perfectly dipolar directions, this configuration cannot be considered as a re-establishment of a normal dipolar configuration, but is a transitional one. A similar transitional configuration has recently been observed during a N-S-N excursion recorded in a lava sequence in the Southern Hemisphere²⁰.

A more stable configuration is observed in the rather long period characterized by intermediate directions and a progressive increase of the field intensity. The duration of this intermediate configuration (30% of the total duration of the reversal) indicates some degree of stability. The Maherado reversal thus provides additional evidence for the existence of transient intermediate states of the geomagnetic field, characterized by non-negligible intensity values and non-dipolar directions. These intermediate states have only been observed previously in lava sequences^{21,22}.

Very large fluctuations in the field-variation rate, such as those documented by this record, have already been observed elsewhere, particularly at Steens Mountain. In our opinion, however, the much more continuous character of sedimentation, compared with the rate of extrusion of lava flow, allows a more precise description of the time characteristics of the reversal.

Some of these characteristics have a bearing on dynamo models. Indeed, it has already been observed that the occurrence of episodes of very rapid changes of the geomagnetic field supports those models for reversal that involve large changes in the level of turbulence in the core under frozen-flux constraints⁶. The Maherado record yields new evidence that favours this type of model. In particular the very fast N-S onset of the reversal seems incompatible with the time constants for flux diffusion reported in the literature and may therefore be entirely controlled by a frozen-flux dominated reversal mechanism.

The characteristics of this record are different from those of the records obtained in Crete^{23,24} from upper Tortonian

sediments, of similar sedimentary and magnetic mineralogy, suggesting that large changes of the transitional fields may occur over a period of ~6 Myr. If confirmed by the other records at Maherado, this result would be a negative test for a recently proposed model for geomagnetic reversals, which requires that all reversals from the same geographical zone be similar within the lifetime of mantle convection patterns²⁵.

Finally, the Maherado record proves that very fine details of geomagnetic reversals can be recovered from carefully selected sedimentary sequences. We consider these sequences, which combine an appropriate chronological control and continuous recording, to be a specifically well adapted support for this kind of study, in spite of the complicated nature of their magnetization. Sedimentary sequences, in particular, allow the correlation of reversals at different sites: thus, if our correlation of the Maherado sequences to anomaly 5A is correct, the complicated structure of the reversal should be observed elsewhere.

We thank A. Mazaud, J. Tessier, J.-P. Valet and P. Tucholka for their help; Professor G. Backus for a discussion with C.L.; Iannis and Katarina Notis for hospitality. The work was supported by the CEA, the CNRS, the INSU-ATP Noyau and NATO. This is contribution 879 from the CFR.

Received 16 June; accepted 7 September 1987.

1. Clement, B. M. & Kent, D. V. *J. geophys. Res.* **89**, 1049–1058 (1984).
2. Theyer, F., Herrero-Bervera, E., Hsu, V. & Hammond S. R. *J. geophys. Res.* **90**, 1963–1982 (1985).
3. Valet, J.-P., Laj, C. & Tucholka, P. *Nature* **322**, 27–32 (1986).
4. Herrero-Bervera, E. & Theyer, F. *Nature* **322**, 159–162 (1986).
5. Prévot, M., Mankinen, E., Grommé, C. & Coe, R. *Nature* **316**, 230–234 (1985).
6. Prévot, M., Mankinen, E., Coe, R. & Grommé, S. *J. geophys. Res.* **90**, 10417–10448 (1985).
7. Jamet, M. thesis, Univ. Orsay (1982).
8. Laj, C., Jamet, M., Sorel, D. & Valente, J.-P. *Tectonophysics* **86**, 45–67 (1982).
9. Kissel, C. thesis, Univ. Paris XI (1986).
10. Hastra, R. thesis, Univ. Utrecht (1982).
11. King, J. W., Banerjee, S. K., Marvin, J. & Ozdemir, O. *Earth planet. Sci. Lett.* **59**, 404–419 (1982).
12. Lowrie, W. & Alvarez, W. *Geology* **9**, 392–397 (1981).
13. Berggren, W. A., Kent, D., Flynn, J. & Van Couvering, J. *Bull. geol. Soc. Am.* **96**, 1407–1418 (1985).
14. King, R. F. *Mon. Not. R. astr. Soc. Geophys. Suppl.* **7**, 115 (1955).
15. Valet, J.-P. & Laj, C. *Earth planet. Sci. Lett.* **54**, 53–63 (1981).
16. Levi, S. & Banerjee, S. K. *Earth planet. Sci. Lett.* **29**, 219–226 (1976).
17. King, J. W., Banerjee, S. K. & Marvin, J. *J. geophys. Res.* **88**, 5911–5921 (1983).
18. Hoffman, K. A. & Slade, J. *Geophys. Res. Lett.* **13**, 483–486 (1986).
19. Burbank, D., Lund, S., Theyer, F. & McRaney, J. *Eos* **67**, 916 (1986).
20. Roperch, P. & Chauvin, A. *Geophys. Res. Lett.* **14**, 151–154 (1987).
21. Shaw, J. *Geophys. J. R. astr. Soc.* **40**, 345–350 (1975).
22. Hoffman, K. A. *Nature* **320**, 228–232 (1975).
23. Valet, J. P. & Laj, C. *Nature* **311**, 552–555 (1984).
24. Valet, J. P., Laj, C. & Langereis, C. G. *J. geophys. Res.* (in the press).
25. Gubbins, D. *Nature* **326**, 167–169 (1987).

A description of energy conversion in photoelectrochemical solar cells

Stuart Licht*

Department of Materials Research, The Weizmann Institute of Science, Rehovot 76100, Israel

Photoelectrochemical cells (PECs) provide alternatives to conventional solid-state solar cells. Since the discovery in 1976, the *n*-cadmium chalcogenide/aqueous polysulphide photoelectrochemical cell, *n*-CdX/S_j²⁻ PEC, has been the most studied and in some ways the most promising of these liquid solar cells^{1,2}, and can combine photoelectrochemical conversion with electrochemical storage³ in a system competitive in overall efficiency with the best comparable solid-state systems. But despite hundreds of studies, neither the reactive chemical species nor the chemical mechanism of operation is known. Here a complete description of the radiative to electrical energy conversion process in *n*-CdX/S_j²⁻ systems is introduced, accomplished by determining the active chemical species and limiting photoelectrochemical processes. Two distinct chemical processes control energy conversion. Photocurrent is limited by supersulphide adsorption. Maximum photopower utilization is desorption limited and related to the activity of the hydrosulphide, hydroxide and shorter polysulphide species.

Photoelectrochemical studies have often focused on solid-state aspects of the systems, accessing available models for analogous solid-state devices. The polysulphide electrolyte used in the *n*-CdX/S_j²⁻ PEC system, formed by the simple dissolution of a sulphide salt and sulphur in water, is known to be highly complex, containing H⁺ and other cations⁴, OH⁻, H₂S, S₂²⁻, the polysulphides (S_j²⁻ (*j* = 2 to 5)) (refs 5,6), supersulphide⁷ (S₂⁻), thiosulphate (S₂O₃²⁻) and sulphate⁶ (SO₄²⁻). As a result of this complexity the overall mechanism of photoinduced oxidation in the *n*-CdX/S_j²⁻ system has remained unknown, but was generally simplified as



Only recently has the emphasis shifted to a pure solution-chemistry investigation of *n*-CdX/S_j²⁻ PECs, leading to some of the highest solar-to-electrical conversion efficiencies in PECs⁸

and to the demonstration of the long-term stability of thin-film solar cells⁹.

This study presents short-circuit photocurrent, *I*_{sc}, radiative-to-electrical energy conversion efficiency, *η*, and fill factor, *f*_i, reflecting the portion of the attainable power accessed, for a *n*-CdSe_{0.65}Te_{0.35} single-crystal photoelectrode immersed in two series of aqueous caesium polysulphide electrolytes, which are measured using a described experimental procedure^{4,8}. Caesium, compared with other polysulphide electrolytes, improves *n*-CdX/S_j²⁻ PEC characteristics⁴.

Figures 1 and 2 summarize the photoelectrochemical and chemical conditions in *n*-CdX/S_j²⁻ PECs using a series of polysulphide electrolytes with consistent hydrosulphide activities and varying pH. These figures include the calculated complex equilibrium distribution of (poly)sulphide chemical species in these electrolytes. In these solutions, at room temperature, spontaneous polysulphide/thiosulphate disproportionation is not significant. Figure 3 summarizes the photoelectrochemical effects in a second series of polysulphide electrolytes containing a nearly constant environment of tetrasulphide and supersulphide.

As Fig. 1 shows, an electrolyte that tends to shift electrolyte composition away from maximum tetrasulphide and supersulphide activities tends to suppress photocurrent. The close linear proportionality of supersulphide activity with short-circuit photocurrent contrasts strongly with the extreme (exponential) variations of most of the sulphide and polysulphide species in solution in Fig. 2. Depletion of the species to be oxidized below an activity level sufficient to utilize photogenerated holes will reduce *n*-CdX/S_j²⁻ photocurrent. These photocurrent losses will tend to be greatest under conditions of maximum band bending (maximum photocurrents). A chemical species whose activity varies linearly with *I*_{sc} is closely related to the process of photoinduced oxidation in *n*-CdX/S_j²⁻.

The importance of the adsorption process in *n*-CdX/S_j²⁻ PECs has been well documented^{1,2}. Supersulphide, rather than tetrasulphide, is likely to be the direct oxidation participant due to: (1) the more proportional relation of *I*_{sc} to supersulphide activity, rather than tetrasulphide activity (see Fig. 1); (2) the higher probability of a one-hole, rather than a two-hole, oxidative process; and (3) comparison of the known temperature dependence of the species distribution in the electrolyte with *I*_{sc}. *n*-CdX/S_j²⁻ *I*_{sc} increases smoothly with temperature¹⁰. Supersulphide displays this smooth increase in activity with temperature⁷, whereas the polysulphide formation equilibrium constants are temperature-insensitive⁵.

* Present address: Department of Chemistry, Room 6-427, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

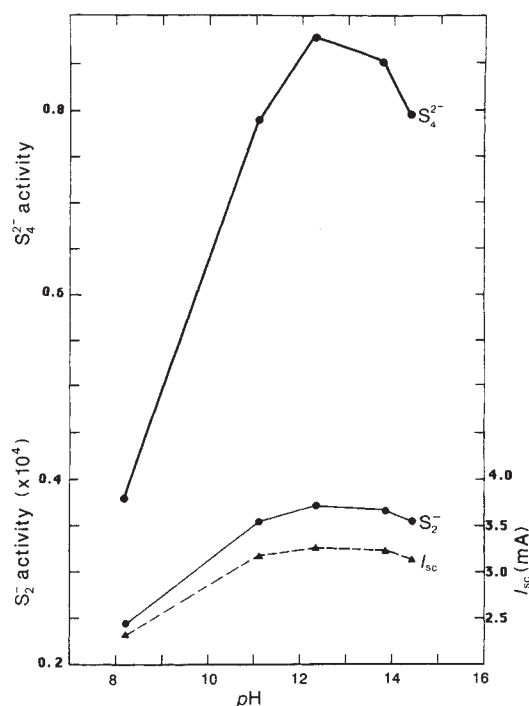
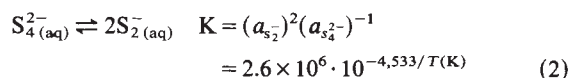


Fig. 1 The variation of supersulphide and tetrasulphide activities and the short-circuit photocurrent I_{sc} , for a 0.15-cm^2 $n\text{-CdSe}_{0.65}\text{Te}_{0.35}$ single crystal immersed in an aqueous electrolyte containing Cs_2S (1.0 mol per kg H_2O), CsHS (0.8 mol per kg H_2O) and sulphur (3 mol per kg H_2O) at various pH values, under an incident filtered tungsten-halogen illumination of 75 mW cm^{-2} . The crystal was mounted with the (0001) plane exposed, and pretreated with mechanical polishing, chemical etching in aqua regia, rinsing in polysulphide and cyanide solutions to remove excess selenium or tellurium, and photoelectrochemical etching in HCl. A 665-nm cutoff filter eliminated differences in light absorption of polysulphide solutions at shorter wavelengths. The pH was varied from an equilibrium value of 12.35 by addition of either H_2S or CsOH to the solution. Chemical activities were calculated by using equilibrium constants^{5,7} at 298 K, in a prescribed manner⁶.

Equilibrium concentrations of the supersulphide species are small, $\sim 10^{-5}$ mol kg^{-1} at room temperature, and insufficient to maintain mass transport at the high current densities measured. However, sufficient tetrasulphide will maintain supersulphide activity, preventing supersulphide depletion at the electrode surface and facilitating the adsorption and oxidation of supersulphide on the surface, M:



In Fig. 3 the nearly constant I_{sc} is again consistent with a constant S_2^{2-} activity, as in equations (3, 4).

Other data appear consistent with the process proposed in equations (2-4). Tetrasulphide was an important participant in $n\text{-CdX}/S_j^{2-}(\text{aq})$ oxidative processes¹¹⁻¹³. Loss of polysulphide in solution by disproportionation resulted in PECs of lowered efficiency^{6,9}. Hamilton and Woods¹⁴ provided evidence for an adsorbed species containing two sulphur atoms, tentatively identified as disulphide, S_2^{2-} , as active in the electrochemical oxidation of sulphide on gold. Lessner *et al.*¹⁵ suggested that supersulphide was a direct participant in the oxidation of polysulphide on cobalt, but their analysis incorrectly excluded adsorptive processes from their model¹⁶.

A significant increase in f_i to >0.5 , parallels increases in

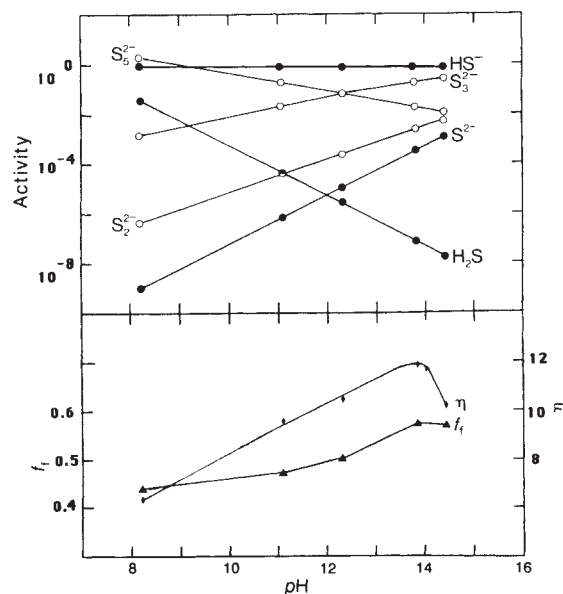
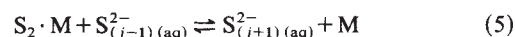


Fig. 2 The variation of fill factor, f_i , solar-electrical conversion efficiency, η , and the distribution of chemical species for a 0.15-cm^2 $n\text{-CdSe}_{0.65}\text{Te}_{0.35}$ single crystal prepared and measured under the conditions described in the legend to Fig. 1. The polysulphide data are shown as open circles, and the hydrogen sulphide, hydrosulphide and sulphide data are shown as solid circles.

hydroxide, disulphide and trisulphide activities (Fig. 2) and hydroxide and hydrosulphide activity (Fig. 3). The depletion of supersulphide activity has a primary effect on photocurrent (Fig. 1) with no relation to the diminishing f_i seen in Figs 2 and 3. Large f_i are maintained only at hydrosulphide and hydroxide activities >0.5 .

At maximum PEC power, band bending is intermediate between the flat band potential and the strong conduction and valence-band bending occurring at I_{sc} . The electrical gradient accelerating holes to the electrode surface is smaller at maximum power than at I_{sc} . As Fig. 3 shows, supersulphide activities are consistently sufficient to maintain high photocurrents. Despite this availability of reductant, the less-reactive holes occurring at maximum power will have a lower probability of favourable (non-corrosive) reaction if the oxidation product tends to accumulate at the electrode surface. As shown, hydroxide and hydrosulphide depletions have a strong effect on f_i , measured at this point of maximum power. This is symptomatic of the different mechanistic roles of supersulphide compared with hydroxide, hydrosulphide and the shorter polysulphide species. The latter species are not directly connected with photocurrent magnitude, but are related to a desorption-dissolution role of these species in the overall oxidative process, according to:



The various polysulphides, $S_{(j-1)}^{2-}(\text{aq})$, will tend to dissolve sulphur (equation (5)). However, a build-up in concentration of the longer polysulphide species will tend to suppress the rate of sulphur dissolution from the surface. This is minimized by the re-equilibration to shorter polysulphide species through equation (6). The sensitive role of hydroxide should be noted. At $\text{pH} < 13$ hydroxide depletion retards sulphur dissolution, and f_i is diminished (Fig. 2). But at $\text{pH} > 14$, tetrasulphide decreases (Fig. 1) as di- and trisulphide activities dominate¹¹⁻¹³. This can impose mass-transport limitations on the photocurrent (equations (2-4)).

Cahen *et al.*¹⁷ observed that aqueous electrolytes, which enhance the dissolution of elemental sulphur, enhance photo-

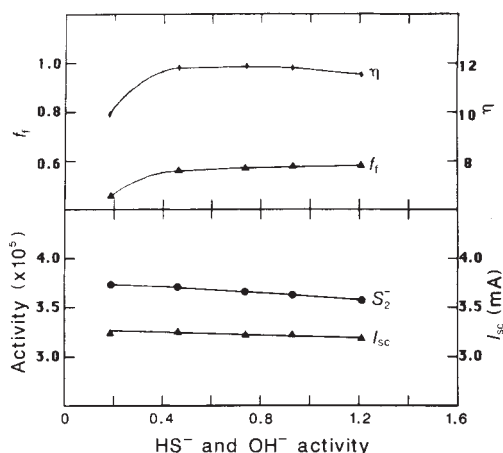


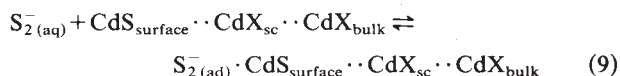
Fig. 3 The short-circuit photocurrent, I_{sc} , solar-electrical conversion efficiency, η , and fill factor, f_i , compared with supersulphide and hydrosulphide (equal to hydroxide) activities, for a 0.15-cm^2 $n\text{-CdSe}_{0.65}\text{Te}_{0.35}$ single crystal immersed in aqueous electrolytes containing sulphur (3 mol per kg H_2O) and Cs_2S (1.0–2.3 mol per kg H_2O), under an incident illumination of 75 mW cm^{-2} . Further measurement and calculation details are described in the legend to Fig. 1.

electrochemical effects in $\text{CdX}/\text{S}_{j(\text{aq})}^{2-}$ PECs. Moreover, the dissolution of an activated sulphur was related to the presence of hydroxide and hydrosulphide¹⁸ although inconsistencies arose in the analysis of the role of hydroxide and sulphide^{19,20}.

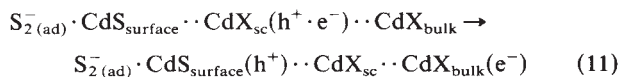
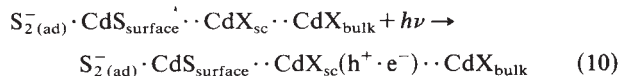
Now equations (2–6) are incorporated into an overall mechanism for the action of $\text{CdX}/\text{S}_{j(\text{aq})}^{2-}$ PECs. In both dark and photoactive environments $\text{Cd}(\text{S}_{x\text{S}}, \text{Se}_y, \text{Te}_z)$ materials, when immersed in polysulphide electrolytes, exchange with sulphur at the surface^{1,2}



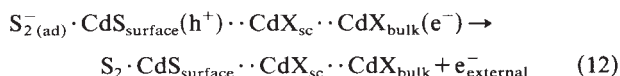
The resultant CdS surface provides stable sites for photooxidation, and must act as a window for photon absorption in the band-gap optimized space-charge region of the semiconductor, CdX_{sc} . Illumination results in useful electrical work and/or chemical fuel production, described here by the following primary process, including equilibration and adsorption of supersulphide,



photoexcitation and separation of electron-hole pair, within the semiconductor,



supersulphide oxidation and external work,

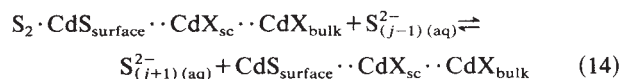


charge balance by reduction at a counter electrode involving either polysulphide reduction, a chemical product or fuel production²¹, or electrochemical storage^{3,22},

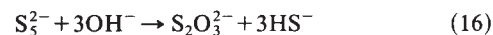


and dissolution of sulphur enhanced by re-equilibration to

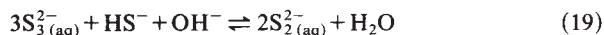
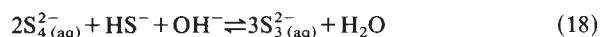
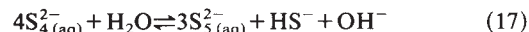
shorter polysulphides,



Processes deleterious to continued photooxidation can be consistently described as adverse to the overall mechanism for the action of $\text{CdX}/\text{S}_{j(\text{aq})}^{2-}$ PECs presented in equations (7–15). These include polysulphide disproportionation⁶, which will remove polysulphide from solution,



a shift of the polysulphide equilibrium away from tetrasulphide^{11–13},



or the enhancement of supersulphide desorption,



electron-hole pair recombination²³,

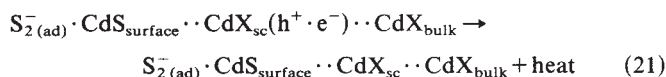
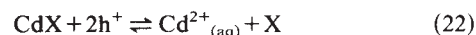
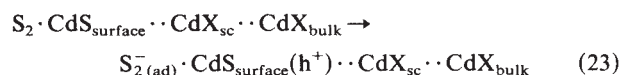


photo-induced corrosion of the semiconductor²⁴,



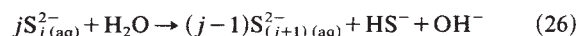
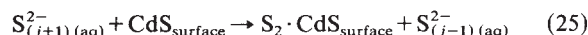
increased dark reactions, indicative of reduction at the semiconductor surface,



or reduction at an alternative material, M, with access to both semiconductor and electrolyte,



and competition to sulphur dissolution, related to depletion of shorter polysulphide species,



This study emphasizes the need for an understanding of solution-chemistry effects in interpreting photoelectrochemical systems. Determination of the reactive chemical species and relevant chemical steps for energy conversion in $n\text{-CdX}/\text{S}_{j(\text{aq})}^{2-}$ PECs, demonstrated here, can be applied to create more accurate modelling of these and other liquid solar cells^{20,23,25–27}.

I thank R. Tenne, Weizmann Institute, Rehovot, Israel, and R. Triboulet and C. Levy-Clement, Laboratoire de Physique de Solides, Meudon, France, for providing crystals.

Received 26 May; accepted 20 August 1987.

- Hodes, G. in *Energy Resources through Photochemistry and Catalysis* (ed. Gratzel, M.) 421–465 (Academic, New York, 1983).
- Hodes, G., Fonash, S. J., Heller, A. & Miller, B. in *Advances in Electrochemistry and Electrochemical Engineering* Vol. 13 (ed. Gerischer, H.) 113B–158B (Wiley, New York, 1985).
- Licht, S., Hodes, G., Tenne, R. & Manassen, J. *Nature* **326**, 863–864 (1987).
- Licht, S., Tenne, R., Flaisher, H. & Manassen, J. *J. electrochem. Soc.* **133**, 52–59 (1986).
- Giggenbach, W. *Inorg. Chem.* **13**, 1724–1730 (1974).
- Licht, S., Hodes, G. & Manassen, J. *Inorg. Chem.* **25**, 2486–2489 (1986).
- Giggenbach, W. *Inorg. Chem.* **10**, 1306–1308 (1971).
- Licht, S. *et al. Appl. Phys. Lett.* **46**, 608–610 (1985).
- Licht, S. *J. phys. Chem.* **90**, 1096–1099 (1986).
- Muller, N. & Cahen, D. *Solar Cells* **9**, 229–245 (1983).

11. Licht, S. & Manassen, J. *J. electrochem. Soc.* **132**, 1076–1081 (1985).
12. Licht, S., Manassen, J. & Hodes, G. *J. electrochem. Soc.* **133**, 272–277 (1986).
13. Licht, S. & Manassen, J. *J. electrochem. Soc.* **133**, 277–280 (1986).
14. Hamilton, I. C. & Woods, R. *J. appl. Electrochem.* **13**, 783–794 (1983).
15. Lessner, P., Winnick, J., McLarnon, F. R. & Cairns, E. J. *J. electrochem. Soc.* **133**, 2517–2522 (1986).
16. Licht, S. *J. electrochem. Soc.* **134**, (in the press).
17. Lando, D., Manassen, J., Hodes, G. & Cahen, D. *J. Am. chem. Soc.* **101**, 3969–3971 (1979).
18. Frese, K. W. Jr & Canfield, D. G. *J. electrochem. Soc.* **131**, 2614–2618 (1984).
19. Licht, S. *J. electrochem. Soc.* **132**, 2801–2802 (1985).
20. Licht, S. & Marcu, V. *J. electroanal. Chem.* **210**, 197–204 (1982).
21. Smotkin, E. S. *et al. J. phys. Chem.* **91**, 6–8 (1987).
22. Bratin, P. & Tomkiewicz, M. *J. electrochem. Soc.* **129**, 2469–2473 (1982).
23. Flaisher, H. & Tenne, R. *J. phys. Chem.* **87**, 3061–3068 (1983).
24. Marcu, V., Tenne, R. & Rubinstein, I. *J. electrochem. Soc.* **133**, 1143–1148 (1986).
25. Orazem, M. E. & Newman, J. *J. electrochem. Soc.* **131**, 2569–2574 (1984).
26. Orazem, M. E. & Newman, J. *J. electrochem. Soc.* **131**, 2574–2582 (1984).
27. Benito, R. M. & Nozik, A. J. *J. phys. Chem.* **89**, 3429–3434 (1985).

Mössbauer imaging

Stephen J. Norton

National Bureau of Standards, Gaithersburg, Maryland 20899, USA

Recoilless γ -ray resonance, or the Mössbauer effect, is a well-established spectroscopic tool in materials science¹. Mössbauer spectroscopy shares some of the basic characteristics of NMR (nuclear magnetic resonance) spectroscopy because both rely on nuclear resonance phenomena. A significant recent advance in the latter field is NMR imaging in biomedicine. The possibility of imaging in NMR arises because the frequency of the nuclear resonance is proportional to the local magnetic-field intensity, and, by imposing a magnetic-field gradient on the spin system, different frequency components in the spectroscopic signal are known to originate at different spatial locations within the sample. In this way, frequency information is translated into spatial information and a map of spin density may be recovered by using well-known image reconstruction algorithms. Here we propose the concept of Mössbauer imaging. In Mössbauer spectroscopy the relative velocity between the γ -ray source and the absorbing object is analogous to the NMR magnetic field. This suggests that an analogous approach to Mössbauer imaging is possible by suitably imposing a velocity gradient on the absorbing object. This can be achieved simply by rotating the absorbing object relative to the source, generating lines of constant Doppler shift (or equivalently, of constant γ -ray energy) across the object. From such measurements, a spatial map of the γ -ray absorption coefficient can be tomographically reconstructed. Spatial resolution is related to the rate of rotation of the absorber, but ultimately is signal-to-noise limited. Although there are no foreseeable applications of Mössbauer imaging in medicine, a variety of applications in materials science are expected. An experimental demonstration of the technique has been performed².

A typical Mössbauer experimental arrangement comprises a γ -ray source, an absorber of the incident γ -rays, and a detector (Fig. 1). In a transmission experiment, the detector counts the number of γ -rays passing through the absorber, and in a scattering experiment, the detector may count the γ -rays re-emitted after resonant absorption. The resonance linewidths of Mössbauer absorption are characteristically very narrow (relative to the γ -ray energies). The possibility of resonance arises due to recoilless emission and absorption of the γ -rays in nuclei embedded in a solid lattice. This resonant absorption can be 'tuned' merely by moving the source at a small velocity relative to the absorber, which imparts some kinetic energy, or Doppler shift, to the γ -ray. Changes in source-absorber speeds as small as a few millimetres per second are frequently sufficient to 'pass through' a resonance line. In a typical experiment, a map of the absorption spectrum of the material is generated by moving the source relative to the absorber and counting the transmitted γ -rays (or re-emitted γ -rays in a scattering experiment) as a function of the relative source-absorber velocity.

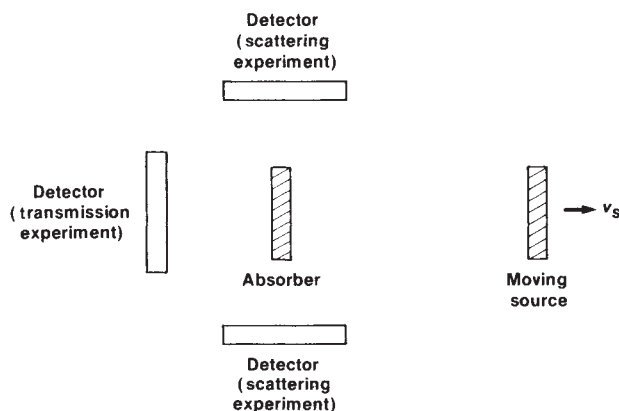


Fig. 1 Conventional Mössbauer experiment.

An image may be viewed as some physical parameter of interest displayed as a function of position, that is, a picture of the parameter as it is distributed in space. In a Mössbauer imaging experiment, the simplest example of such a parameter is the Mössbauer absorption coefficient; the aim is to reconstruct and display its spatial distribution. A conventional experiment measures only a bulk average of the resonant absorption coefficient over the absorbing sample; thus spatial inhomogeneities go undetected in the averaging process. A Mössbauer imaging experiment, however, would reveal variations in the two-dimensional distribution of the absorption coefficient. In a slightly more elaborate imaging experiment, true spectroscopic information can be recovered as a function of position³. The ability to perform Mössbauer spectroscopy in an imaging mode, that is, to reconstruct the Mössbauer spectrum as a function of position within the sample, rather than in bulk, should be of significant value in the study of materials with spatially varying properties.

In principle, a crude method of imaging the absorption coefficient distribution would consist of collimating the γ -radiation by a suitable arrangement of slits (Fig. 2) and directing the collimated beam at the absorber over a wide range of incident angles. This approach is entirely analogous to X-ray computerized tomography (CT) in a transmission experiment, and similar to single-photon emission CT in a scattering experiment. As discussed below, however, we can achieve the same result without slits by rotating the object and simultaneously translating the source. Moreover, this approach has a significant advantage over methods that use slits when the source is spatially extended. The ability to use an extended source is important because of the resulting greater incident γ -ray flux and corresponding signal-to-noise ratio enhancement. Finally, an advantage of the proposed imaging approach over conventional (non-resonant) X-ray imaging techniques is that it is exceedingly isotope-specific, that is, the γ -ray resonant absorption cross section of a particular Mössbauer element is many orders of magnitude larger than the non-resonant X-ray absorption cross section.

To explain how a Mössbauer imaging system might arise, it is helpful to consider the analogy of NMR imaging. In NMR spectroscopy, the resonance frequency (the precessional frequency of the nuclear spins) is proportional to the local

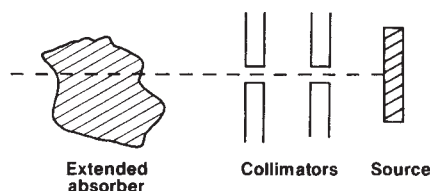


Fig. 2 Collimating the γ -ray source.

magnetic-field strength. By sweeping the strength of an applied magnetic field in the presence of an r.f. excitation field, resonant absorption may be mapped as a function of frequency (or applied field strength). In an NMR imaging experiment, magnetic field gradients are applied that create lines (in two dimensions) or surfaces (in three dimensions) of constant field strength, and hence lines or surfaces of constant resonance frequency. By analogy, by imposing a velocity gradient on a spatially extended Mössbauer absorber, similar lines (in two dimensions) or surfaces (in three dimensions) of constant resonant absorption should be measurable. In an equivalent interpretation, one wishes to move the absorber and source relative to each other so as to create lines or surfaces of constant γ -ray Doppler shift.

This result can be achieved simply by rotating the absorber with constant angular speed while translating the source at constant velocity relative to the absorber. (We shall assume constant source velocity for simplicity; the method, however, can be readily generalized to the case of accelerated source motion.) Figure 3a shows the absorber rotating about the z -axis (into the page) and the source moving at constant speed v_s away from the absorber along the x -axis. We assume, for the moment, that the source-absorber distance is large so that the velocity of the impinging γ -rays may be regarded as parallel to the x -axis. We also assume a two-dimensional absorbing specimen, where the absorber lies in the plane of the incident radiation. This 'edge-on' geometry is unusual in Mössbauer spectroscopy, but is helpful in visualizing the imaging process. Elementary geometric considerations show that the x -component of the instantaneous velocity of rotation of the absorber, rotating about the z -axis, is constant along any line parallel to the x -axis, and the magnitude of this velocity is just $v_x = R\Omega$, where R is the radial distance of this line from the centre of rotation and Ω is the angular speed of rotation (Fig. 3a). This can be seen by writing one point on the rotating absorber in polar coordinates: $x = r \sin \Omega t$ and $y = R - r \cos \Omega t$; hence, $v_x = \dot{x} = R\Omega$. Thus, merely by rotating the object, one creates a linear gradient in the x -component (v_x) of the velocity of rotation, where the direction of the gradient is in the positive y -direction (when rotating clockwise) and the magnitude of the gradient is $\Omega \text{ s}^{-1}$. As the absorber rotates clockwise and the source moves in the positive x -direction at constant velocity v_s , resonant absorption will take place only along the line for which $v_x = v_s$ holds, or when $R = v_s/\Omega$. If the number of detected γ -rays is recorded as a function of time, it is clear that line integrals of the absorption coefficient are measured, also as a function of time, over all straight lines through the absorber tangent to the circle of radius $R = v_s/\Omega$. Hence, by varying v_s or Ω , one can, in principle, generate a complete set of line integrals of the absorption coefficient and thus acquire enough data to reconstruct its distribution by using well-known tomographic algorithms^{3,4}. If the γ -ray attenuation is significant because of resonant absorption or other mechanisms, so that the number of resonance events is not uniformly distributed along the line, then iterative tomographic algorithms designed to take attenuation into account should be used (for example, ones similar to those used in single-photon emission CT)⁴.

Spatial resolution in this technique is determined by the ratio of the natural γ -ray linewidth (in units of velocity) to the velocity gradient Ω , that is, if the Mössbauer linewidth is Δv , the spatial resolution is $\Delta v/\Omega$. Thus, it appears that one could achieve arbitrarily high resolution merely by increasing the rate of rotation. As an example, for a typical line width of $\Delta v = 1 \text{ mm s}^{-1}$, a rate of rotation of several hundred revolutions per second implies, in principle, a spatial resolution in the submicrometre range. In practice, the resolution will be signal-to-noise limited and such resolution would require exceedingly long integration times because of the relatively low intensity of available Mössbauer sources.

The situation can be noticeably improved by moving the

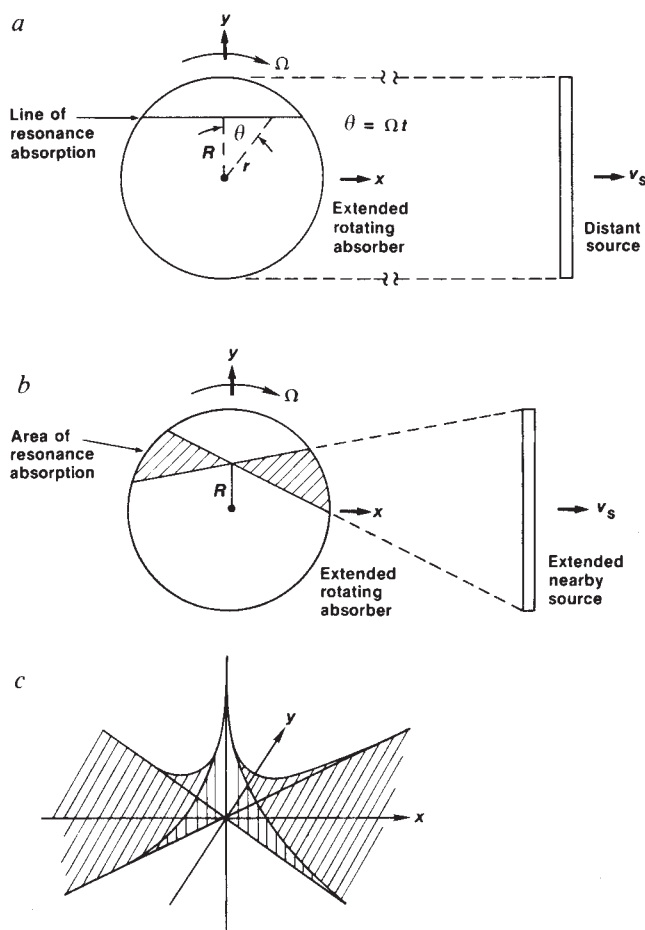


Fig. 3 *a*, A single line exists along which resonant absorption takes place when the γ -rays are incident parallel to the x -axis (Conditions for resonance absorption: $R\Omega = v_s$). *b*, For a nearby extended source, this line broadens into the sector as shown. *c*, Probability distribution of absorption per unit area over the region of resonance absorption in (*b*). The probability of absorption per unit area peaks at $(0, v_s/\Omega)$. The distribution is narrow in the y -direction but falls off as $1/|x|$ in the x -direction.

source nearer to the absorber, in which case the source must now be regarded as spatially extended in the y -direction. Figure 3b shows such a source moving in the x -direction at constant speed v_s . In this geometry, the region within the (two-dimensional) absorber (rotating at angular speed Ω) over which resonant absorption will take place can easily be shown to be a sector, as indicated in Fig. 3b, where the boundaries of the sector are straight lines passing through the end points of the extended source and intersecting at the point $(x=0, y=v_s/\Omega)$ within the absorber. Neglecting, for simplicity, attenuation due to absorption (that is, assuming most of the γ -rays are not captured along their path through the absorber), the probability distribution of absorption per unit area is simply proportional to the inverse of the sector area shown in Fig. 3b, that is, the distribution peaks at the point $(0, v_s/\Omega)$ and falls off as $1/|x|$ in the x -direction. Through the peak in the y -direction, the fall-off is much more rapid; the width along the y -axis is again $\Delta v/\Omega$, where Δv is the natural linewidth of the Mössbauer element. The probability distribution of absorption events is shown in Fig. 3c. This peaked distribution is significant because, merely by varying v_s and counting absorption events as a function of time, it should be possible to generate a blurred image without additional computer processing. It is important to appreciate that the above distribution (Fig. 3c), with the narrow peak of width $\Delta v/\Omega$ along the y -direction, cannot be achieved with any arrangement of collimating slits.

The imaging process is particularly simple to analyse when the Mössbauer element has a single dominating absorption line. Many elements, however, exhibit more complex spectra comprising multiple lines. A particularly complex or extended spectrum may complicate the imaging process. An analogous problem arises in NMR imaging, because in the process of imaging, spectral (frequency) information is translated into spatial information; as a result, true spectroscopic information gets 'folded' into the spatial response of the system. This presents no problem if the gradient (magnetic, or velocity, in NMR, or Mössbauer, respectively) is sufficiently large so that the spectroscopic bandwidth is confined to a single spatial resolution element in the image. In a more elaborate version of the above Mössbauer imaging experiment, one can recover the actual extended Mössbauer spectrum at each point in the image (that is, perform spectroscopy in an imaging mode). In this process, one reconstructs a three-dimensional function with two spatial dimensions and one spectral dimension. This result is achieved by counting the resonantly absorbed γ -rays as a function of three degrees of freedom in the measurement: Ω , v_s and time (or, equivalently, the instantaneous position of the rotating absorber during the measurement)³. A somewhat analogous approach has also been proposed in NMR chemical-shift imaging, where the aim is to recover the NMR spectrum as a function of position. This is accomplished by introducing an additional degree of freedom into the NMR measurement, for example, by varying the magnitude of the gradient rather than merely its direction⁵.

Applications of Mössbauer imaging to the interior of objects appear to be limited to lighter materials, or small objects, that permit sufficient γ -ray penetration. One possibility is the imaging of composite materials, in which one wishes to image the distribution of a Mössbauer element embedded in a lighter matrix. Another possibility is surface imaging, whereby the surface of a flat absorber is rotated about an axis normal to the surface and a one-dimensional extended source is displaced a small amount from the plane of the surface. A number of specific metallurgical applications are also proposed in the following paper presenting experimental results².

Received 8 June, accepted 25 August 1987.

1. Danon, J. *Lectures on the Mössbauer Effect* (Gordon and Breach, New York, 1968).
2. Atzmony, U., Norton, S. J., Swartzendruber, L. J. & Bennett, L. H. *Nature* **330**, 153–154 (1987).
3. Norton, S. J. *J. Res. natn. Bur. Stand.* Vol. 92, 325–334 (1987).
4. *IEEE Trans. nucl. Sci.* NS-21 (1974) (Special issue on image reconstruction); *Comput. Biol. Med.* **6** (1976) (Special issue on image reconstruction).
5. Brown, T. R., Kincaid, B. M. & Ugurbil, K. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3523–3525 (1982).

Mössbauer imaging: experimental result

U. Atzmony*, S. J. Norton, L. J. Swartzendruber & L. H. Bennett

National Bureau of Standards, Gaithersburg, Maryland 20899, USA

In the previous paper¹, the concept of Mössbauer imaging was proposed. Here we report the first experimental demonstration of this concept. For simplicity, a one-dimensional imaging experiment is described; however, the fundamental imaging principle thus demonstrated generalizes, in a straightforward way, to higher dimensions. Finally, we speculate on some possible applications of this technique in materials science.

The fundamental concept of Mössbauer imaging is that, by imposing a velocity gradient on the absorber, different spatial locations in the absorber will experience different velocities relative to the source; in this way, the resonant absorption distribution will be spatially resolved. The spatial distribution can then be reconstructed by standard tomographic imaging algorithms¹.

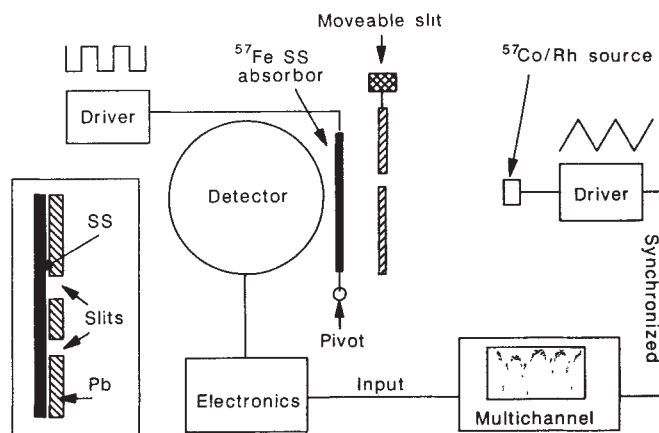


Fig. 1 Schematic diagram of the experimental arrangement for one-dimensional Mössbauer imaging. For calibration, a stainless-steel foil was used as an absorber and spectra were taken with a movable slit. The inset shows the sample used to obtain the spectrum of Fig. 2.

In one dimension, the imaging principle can be most easily demonstrated as no complicated tomographic reconstruction is necessary; the image, that is, the resonant absorption coefficient displayed as a function of position, is generated directly. As originally described¹, a two-dimensional imaging experiment involves rotating the absorber about a fixed centre. In a one-dimensional experiment, where the source moves with a constant acceleration in the x -direction, it suffices merely to pivot the absorber periodically through a small angle about the vertical (as shown in Fig. 1). With the pivoting end of the extended absorber fixed, the instantaneous velocity along the length of the absorber will vary linearly from zero at the pivot position to a maximum value at the opposite end. Thus, if $v_x(t, y)$ is the x -component of the absorber velocity at the point y at time t , then

$$v_x(t, y) = (y/y_{\max})v_x(t, y_{\max})$$

where $v(t, y_{\max})$ is the velocity of the end of the absorber. Assuming a single-line absorber, the condition for resonant absorption is then simply

$$v_x(t, y) = v_s(t) + v_i$$

where v_s is the source velocity and v_i is the isomer shift.

In the one-dimensional imaging experiment reported here, the high end of the absorber (Fig. 1) was driven with a square-wave velocity, where the magnitude of the velocity alternated between $+6$ and -6 mm s⁻¹, and the frequency was 1.2 Hz. The source was driven with constant acceleration (sawtooth velocity waveform) over a range of velocities of ± 6 mm s⁻¹ at a frequency of 1.6 Hz. A multi-channel analyser, synchronized to the source-motion, was used to accumulate spectra. It is not necessary to synchronize the motion of absorber to the motion of the source. The spectrum obtained with this arrangement consists of four identical spectra; these are reduced to one spectrum by appropriate folding.

The absorber was a 10 mm by 10 mm 310 stainless steel foil, 12 μ m thick and enriched to 91% ⁵⁷Fe. The source was approximately 50 mCi ⁵⁷Co in Rh. Two narrow slits, 4.4 mm apart, in a lead foil shield were used to simulate an absorber with iron present over two narrow bands. Figure 2 shows the resulting spectrum, accumulated in ~ 2 h. The positions of the dips show accurately the positions of the slits and the magnitudes of the dips show accurately the relative width of the two slits. (The length scale in Fig. 2 was calibrated with a movable slit.) Of course, it would have been possible to obtain the same information in the spectrum of Fig. 2 without imaging; by using a narrow slit, a large number of spectra could be measured as the slit is repositioned from the top to the bottom of the absorber, but

* Permanent address: Nuclear Research Center-Negev, Beer-Sheva, Israel.

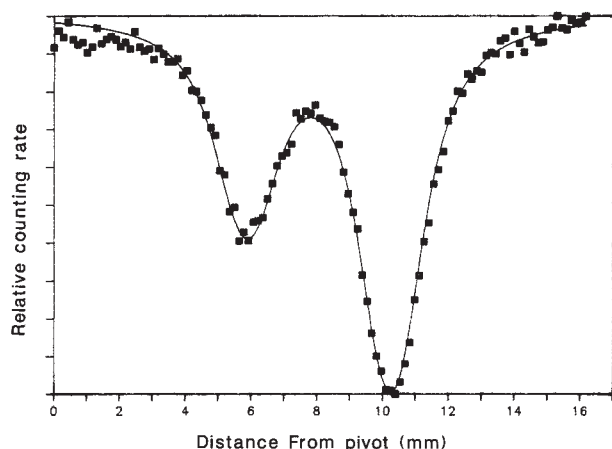


Fig. 2 One-dimensional Mössbauer image showing the distribution of stainless steel across a sample. The sample contained two narrow bands of stainless steel as seen in Fig. 1.

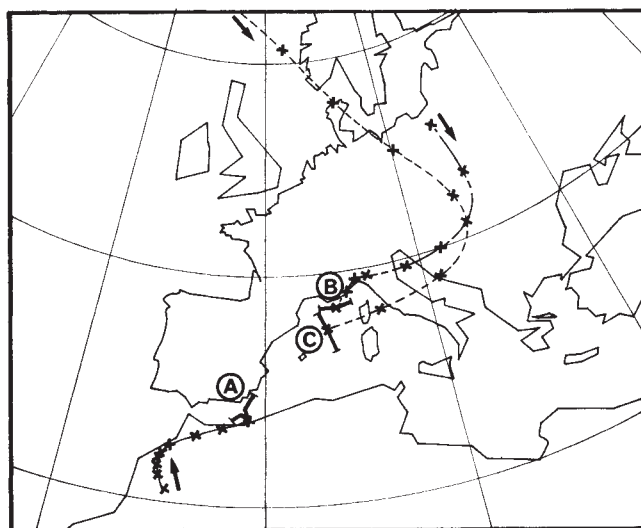


Fig. 1 Sampling location and air trajectories of Mediterranean aerosol samples (A: October 17–18, 1981; B: October 25, 1983; C: October 23–24, 1983). The travel time of the air between two successive crosses is 12 h. The air trajectories are calculated over a period of 4 d, arriving at the 925-hPa level at the mid-point of each sampling time. They have been computed from the analysed wind fields of the European Centre for Medium-Range Weather Forecasting of Reading, UK^{18,19}. The calculations have also given the altitude of air transport¹¹, based on an approximation of the vertical component of the wind.

the time required to accumulate the equivalent resolution would be much longer than in the imaging mode.

It is evident from Fig. 2 that one-dimensional imaging of a Mössbauer absorber has been achieved. Even such a simple imaging technique could be useful in materials research or production. As an example, longitudinal variations in composition or in internal stresses of rapidly quenched ribbons of amorphous alloys containing iron could be monitored. Although practical limitations in our crude electronic and mechanical arrangement restrict the resolution achieved here, in principle the resolution can be greatly enhanced by increasing the velocities at which the source and absorber are driven. Sufficiently increased velocities should also make possible the imaging of ferromagnetic materials.

It has been described in the previous paper¹ how an extension to two-dimensional imaging may be achieved. Such an extension would be of greater practical value in the inspection of samples in the shape of foils. Extension to Mössbauer scattering is more involved, but could increase the non-destructive testing applications of the method.

Received 8 June; accepted 25 August 1987.

1. Norton, S. J. *Nature* **330**, 151–153 (1987).

Stable lead isotope tracers of air mass trajectories in the Mediterranean region

Hal Maring*, Dorothy M. Settle*,
Patrick Buat-Ménard†, François Dulac†
& Clair C. Patterson*

* California Institute of Technology, Division of Geological and Planetary Sciences, Pasadena, California 91125, USA

† Centre des Faibles Radioactivités, Laboratoire Mixte CNRS-CEA, BP 1, F-91190, Gif-sur-Yvette, France

Stable lead isotopes in industrial lead-rich aerosols can be used as tracers of atmospheric air masses and to identify the geographical location of atmospheric lead emission sources. Industrial lead emissions outweigh natural emissions by two orders of magnitude and therefore determine the isotopic composition of atmospheric lead¹. We report preliminary data which indicate that stable lead isotopes are effective in identifying source regions for aerosols from a variety of industrial lead emission sources in complicated atmospheric systems such as the Mediterranean region.

Isotopic compositions of lead-rich aerosols in the atmosphere are different in various geographical regions because the geological origin of lead ores supplied to and used by different nations varies. Variations in lead isotope abundance between ore bodies of different age and location result from the different accumulation of daughter lead, generated by radioactive decay from uranium and thorium that has been associated with the lead during geological time, and depends on how long lead and its radioactive parents were together in the mantle and crust before the lead was segregated into an ore body.

There are four stable lead isotopes, but the ²⁰⁶Pb/²⁰⁷Pb ratio is generally used for tracer purposes because it can be most accurately measured. This ratio is ~1.10 in some 1,000-Myr-old ore deposits, ~1.15 in 500-Myr-old ores, and ~1.19 in 30-Myr-old ores². These changes conform to the isolation of lead ores at progressively later times from a source rock system containing relatively fixed ratios of trace quantities of uranium and lead. Some lead ore bodies are derived from source rocks which contain unusually large concentrations of uranium, as in the case for lead ores mined from Missouri in USA, which are characterized by ²⁰⁶Pb/²⁰⁷Pb ratios of ~1.30, even though the deposits are only a few hundred million years old². Some lead ores are derived from extremely old source rocks, such as those from Broken-Hill, Australia, which are ~2,000-Myr-old and are characterized by ratios of ~1.05.

Lead isotope tracers have been used to prove that the origins of aerosols emitted from land sources and then transported within the North Pacific Easterlies vary systematically with season³. The ²⁰⁶Pb/²⁰⁷Pb ratio in aerosols collected on filters during April 1979 at Enewetak was 1.170, and in August at the same location it changed to 1.205. The estimated average ²⁰⁶Pb/²⁰⁷Pb of industrial lead emissions from Asia/Japan at this time ranged from 1.153 to 1.165, whereas the estimated average ratio of lead emissions from western United States/Mexico ranged from 1.230 to 1.187. These data show that air masses crossing Enewetak in April came from north-west Asia, and that those in August came from United States/Mexico regions east of Enewetak. Mineralogical tracer studies of these air masses

Table 1 Stable isotopic compositions of lead from ore deposit in Morocco and eastern Europe and from Italian and suburban Paris air

Location	$^{206}\text{Pb}/^{207}\text{Pb}$ range	$^{206}\text{Pb}/^{208}\text{Pb}$ range	$^{206}\text{Pb}/^{204}\text{Pb}$ range
Suburban Paris air ¹⁴	1.101	0.4601	17.25
Moroccan ores ^{10,15}	1.149–1.186	0.4696–0.4810	17.96–18.69
Yugoslavian ores ^{10,15}	1.193–1.199	0.4788–0.4808	18.75–18.92
Yugoslavian ores ^{9,15}	1.162–1.175	0.4735–0.4777	18.20–18.39
Greek ores ^{10,15}	1.195–1.200	0.4822–0.4854	18.80–18.99
Greek ores ^{9,15}	1.167–1.196	0.4770–0.4831	18.25–18.77
Italian air ¹⁵	1.185	—	—
Calculated average of east Europe, Yugoslavia and Greece	1.186	—	18.62
Calculated average of 1 to 1 mix Italy plus France	1.144	—	17.94

at the same times at Enewetak, also using air filters, gave similar indications^{4,5}. Retrospective air trajectory analysis by meteorologists indicates that easterly air masses crossing Enewetak in April 1979, probably originated from Asia, and in August they probably originated from meso America⁶.

The data presented here indicate that stable lead isotope measurements can be used to identify the origins of aerosols in more complex atmospheric systems, such as the Mediterranean region, which are supplied by a variety of peripheral and distant industrial lead emissions sources⁷. In this study geographical origins of various leads on air filters collected aboard ship in the Mediterranean are assigned by comparisons with isotopic composition of ores which are presumed to provide the lead emitted from a given geographic region. This is done because only two isotopic compositions of lead in urban atmospheres are available to us at present (France and Italy), and furthermore, ore leads are the major sources (lead alkyls), relative to non-ore sources (coal combustion), of pollution lead over Europe⁸. Table 1 lists values for stable lead isotopic compositions of lead from some ore deposits used for industrial supplies in Morocco and in eastern Europe, as well as from ambient air filters operated in suburban Paris and northern Italy. The extremely low $^{206}\text{Pb}/^{207}\text{Pb}$ ratio observed in Parisian air is a reflection of the use of Australian ores as a source of industrial lead in France today. The range of ratios given for Morocco is derived from a number of North African ores. One range of ratios for Yugoslavian ores⁹ originates from a variety of minerals from a single deposit and the other range of values¹⁰ expresses values of different galenas taken from one ore deposit. One range of values given for Greece expresses the variation from less radiogenic ores in northern Greece to more radiogenic ores at Laurion and in the southern islands⁹, and the other range originates from one ore deposit in southern Greece¹⁰.

Bulk atmospheric aerosol samples were collected on board ship during a series of oceanographic cruises in the western Mediterranean Sea by investigators from the Centre des Faibles Radioactivités (Fig. 1). Ultra-clean collection was performed from a bow tower, using wind direction control to avoid contamination from the ship. The aerosols were sampled on 47-mm acid-washed Whatman 41 cellulose filters. Stable lead isotopic analyses were performed at the California Institute of Technology using a thermal ionization source, high-resolution mass spectrometer after acid digestion of the filter and dithizone extraction concentration of the lead³.

The isotopic compositions of lead in aerosol samples collected on two cruises, one in 1981 and one in 1983, were determined, and results of these analyses are listed in Table 2, together with probable regional origins of aerosols in the sample air masses¹¹ (see also Fig. 1). On 17–18 October, 1981, an air filter sample was collected aboard RV *Le Cornide de Savedra* positioned off the North African shore of western Algeria. The atmospheric lead concentration determined by flameless atomic absorption spectroscopy¹¹ was $56 \pm 5 \text{ ng m}^{-3}$. This value is much higher than that generally found over the open North Atlantic, thus suggesting the influence of close land-based sources. The isotopic

composition of lead in this sample is similar to that of Moroccan lead ores. The isotopic compositions of leads in Moroccan urban air are not available to us. If it is assumed that the isotopic composition of lead in Moroccan ores (Table 1) represent that in Moroccan air, then the isotopic composition of lead in this air filter sample suggests that the air did originate from North Africa. Such an interpretation is supported by retrospective air mass trajectory analyses for this date and location. Three-dimensional, four-day back trajectories were calculated for air arriving at three different levels: 925 hPa (trajectory A in Fig. 1), 700 and 500 hPa. They all indicate that the source region of the sampled air mass was only north-west Africa, with no probable contributions from Europe. Further, this air sample is characterized by a high concentration of soil-dust material, as indicated by the concentration of Al ($\sim 700 \text{ ng m}^{-3}$) (ref. 11). This concentration is 3–4 times higher than the geometric mean Al concentration found over the western Mediterranean¹¹, and such high concentrations over this marine environment have in all cases been found to be related to dust transport episodes from North Africa or Sahara^{11,12}.

Two air-filter samples were collected in October 1983 aboard RV *Le Suroît* in the western Mediterranean region, south of Marseilles and west of Corsica. Retrospective trajectories for these two successively sampled air masses show that both passed over the same general regions of eastern and southern Europe before their pathways diverged upon leaving Italy (paths B and C in Fig. 1). After leaving Italy, retrospective air mass trajectory analysis and geochemical data¹¹ indicate that air following pathway B passed directly through the heavily urbanized Marseilles region before travelling south a short distance out to sea, where it was sampled. On the other hand, air following pathway C left the coast of Italy and travelled west a relatively long distance over the sea before it was sampled.

Lead filtered from air which followed path B was found to exhibit an isotopic composition similar to lead alkyls burned in French air (Table 1), which presumably characterizes that in urban air of the Marseilles region. This indicates that lead of a quite different isotopic composition, presumably introduced upstream into the air mass from eastern Europe, and certainly

Table 2 Lead isotopic compositions in aerosol samples and their origins indicated by retrospective air mass trajectory analysis (see Fig. 1)

Sample	Probable regional origin of aerosols	$^{206}\text{Pb}/^{207}\text{Pb}$ *	$^{206}\text{Pb}/^{208}\text{Pb}$ †	$^{206}\text{Pb}/^{204}\text{Pb}$ ‡
A (Oct. 17–18, 1981)	North-west Africa	1.156	0.4757	17.96
B (Oct. 23–24, 1983)	France and Italy	1.149	0.4744	17.82
C (Oct. 25, 1983)	Marseilles	1.118	0.4680	17.27

* Accuracy is ~ 1 part in 1,000; † accuracy is ~ 2 parts in 1,000; ‡ accuracy is ~ 5 parts in 1,000.

introduced from Italy (Table 1), had either been removed for the most part by the time the air had reached Marseilles, or its abundance in air was simply overwhelmed by the introduction of large excesses of Marseilles gasoline exhaust lead. The latter explanation is supported by the very high concentration observed for this sample: $144 \pm 14 \text{ ng m}^{-3}$. Moreover, consideration of the level of air transport¹¹ indicates that this high concentration can be attributed to the long residence time of the air (~ 3 days) in the lower troposphere ($< 1,500 \text{ m}$), where pollutants are emitted.

On the other hand, lead filtered from air which travelled path C was present in low concentration: $6 \pm 1 \text{ ng m}^{-3}$. The consideration of the level of air transport indicates that the sampled air mass had only spent a few hours at low level, over Italy and Corsica. The isotopic composition is similar to that expected of a mixture of equal parts of burned lead alkyls from French air and from Italian air. Even though the total emission flux of gasoline lead from Corsica island is small compared to that of the Italian mainland, the isotope data for lead in this sample indicates that passage over Corsica greatly altered the character of the lead from a presumably Italian gasoline-exhaust isotope type to a mixture which includes another French isotope type of exhaust lead. Alternatively, one cannot exclude that contribution of lead from Marseilles could have occurred, because the trajectory was reconstructed for the mid-point of the sampling period (37.5 hours duration), and therefore may have missed such a contribution introduced near the end of the sampling period¹³.

Both these latter two samples of filtered marine air indicate that care must be taken when generalizing information on an air trajectory to evaluate contributions from local and distant sources to the aerosol burden on a timescale of a day or more. These examples indicate that lead tracer data in meteorologically identified parcels of marine air probably backtrack only to the latest urban terrestrial source. This can serve a useful purpose in answering questions retrospectively which concern the relative importance of different air parcel components which have travelled from geographically different land sources to mix before reaching a marine atmospheric sampling site.

It is important to have available an extensive and reliable database delineating the isotopic compositions of lead in aerosols at various geographic sources to use the lead isotope tracer method most effectively. A collation of existing data in the literature, concerning lead isotope data in aerosols and their provenance, should be augmented by collection and isotope analysis of additional samples which would document recent temporal changes in lead isotopic compositions of aerosols from different geographic regions. This would also contribute to a better understanding of the perturbation of the lead cycle in marine systems which is primarily due to anthropogenic lead inputs of eolian origin^{16,17}.

This work was supported by the NSF and by CNRS, CEA, IFREMER and the Ministère de l'Environnement in France. We are indebted to D. Martin, B. Strauss and the reviewers for constructive comments.

Received 2 March; accepted 16 September 1987.

- Patterson, C. C. & Settle, D. M. *SEAREX Newsletter* 10, No. 1 (1987).
- Doe, R. B. *Lead Isotopes* (Springer, Berlin, 1970).
- Settle, D. M. & Patterson, C. C. *J. geophys. Res.* **87**, 8857–8869 (1982).
- Buat-Ménard, P., Ezat, U. & Gaudichet, A. *Precipitation Scavenging, Dry Deposition and Resuspension* Vol. 2, 1259–1270 (Elsevier, New York, 1983).
- Duce, R. A., Unni, C. K., Ray, B. J., Prospero, J. M. & Merrill, J. T. *Science* **209**, 1522–1524 (1980).
- Merrill, J. T., Bleck, R. & Avila, R. *J. geophys. Res.* **90**, 12927–12936 (1985).
- Buat-Ménard, P. in *Strategies and Advanced Techniques for Marine Pollution Studies: Mediterranean Sea* (eds Giam, C. S. & Dou, H. J. M.) 187–199 (Springer, Berlin, 1986).
- Pacyna, J. *Atmos. Envir.* **18**, 41–50 (1984).
- Settle, D. M. & Patterson, C. C. *Use of Stable Lead Isotopes as Provenance Indicators for Lead Artifacts in Classic Greek and Roman Cultures* (Lecture, National Archeological Museum, Athens, 1972).
- Russel, R. D. & Farguhar, R. M. *Lead Isotopes in Geology* (Wiley, New York, 1960).
- Dulac, F., Buat-Ménard, P., Arnold, M., Ezat, U. & Martin, D. *J. geophys. Res.* **92**, 8437–8453 (1987).

- Chester, R., Sharple, E. J., Sanders, G. S. & Saydam, A. C. *Atmos. Envir.* **18**, 929–935 (1984).
- Dulac, F. thesis, Univ. Paris 7 (1986).
- Elbaz-Poulichet, F., Holliger, P., Huang, W. W. & Martin, J. M. *Nature* **308**, 409–414 (1984).
- Facchetti, S. *et al. Isotopic Lead Experiment*, 114 pp. (Status Report, Roma, Commission of the European Communities, EPA and ILZO, USA and Societa Italiana Additivi 1982).
- Flegal, A. R. & Patterson, C. C. *Earth planet. Sci. Lett.* **64**, 19–32 (1983).
- Veron, A., Lambert, C. E., Isley, A., Lmet, P. & Grousset, F. *Nature* **326**, 278–280 (1987).
- Lorenc, A., Rutherford, I. & Larsen, G. *European Centre for Medium-range Weather Forecast Technical Report No. 6* (Reading, 1977).
- Martin, D., Mithieux, C. & Strauss, B. *Atmos. Envir.* **21**, 41–52 (1987).

Cd/Ca in late Miocene benthic foraminifera and changes in the global organic carbon budget

Margaret Lois Delaney* & Edward A. Boyle†

* Institute of Marine Sciences, University of California, Santa Cruz, California 95064, USA

† Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology E34-258, Cambridge, Massachusetts 02139, USA

The late Miocene carbon shift ($\sim 6.2 \text{ Myr}$)—a $0.5\text{--}1.0\%$ $\delta^{13}\text{C}$ decrease in benthic and planktonic foraminifera—has been ascribed to changes in global inventory, deep-ocean circulation, and/or productivity^{1–16}. Cadmium, $\delta^{13}\text{C}$, and nutrients in the ocean are linked; comparison of $\delta^{13}\text{C}$ and Cd/Ca yields circulation and chemical inventory information not available from either alone^{17,18}. We determined Cd/Ca ratios in late Miocene benthic foraminifera from DSDP Site 289. Results include: (1) late Miocene Pacific Cd/Ca values fall between those of late Quaternary Atlantic and Pacific benthic foraminifera; (2) there are no systematic Cd/Ca offsets between *Cibicides kullenbergi*, *Cibicides wuellerstorfi* and *Uvigerina* spp.; and (3) there is a very slight Cd/Ca change coincident with $\delta^{13}\text{C}$. Cd/Ca, slightly higher in younger, isotopically lighter samples, exhibits a smaller increase than predicted if circulation were the primary cause of the carbon shift. The carbon shift may have been due to a long-term shift in the steady-state carbon isotope input or to a change in the sedimentation of organic carbon relative to calcium carbonate.

The abundance of information about DSDP Site 289 (Leg 30), located on the Ontong-Java Plateau in the western equatorial Pacific ($00^{\circ}29.92'\text{S}$, $158^{\circ}30.69'\text{E}$; 2,224-m water depth), made it ideally suited for the first measurements of Cd/Ca in benthic foraminifera older than the Pleistocene (and for times much longer than the Cd residence time). Site 289 has been studied for foraminiferal biostratigraphy^{19–21}; the historical record of carbon and oxygen isotopes in planktonic and benthic foraminifera^{7,22} and in bulk calcium carbonate²³; Sr/Ca and Mg/Ca in bulk calcium carbonate^{23,24}; interstitial water chemistry, including oxygen and strontium isotopes²³; and Li/Ca, Sr/Ca, Mg/Ca and $^{87}\text{Sr}/^{86}\text{Sr}$ in planktonic foraminifera^{23,25}. The Miocene carbon isotope stratigraphy of Site 289 has been used as a reference section for correlation with other sites^{8,11–13}. The late Miocene carbon shift, the most thoroughly documented and modelled Miocene carbon event, was chosen for this study; the composition of foraminiferal shells of this age at Site 289 do not appear to have been substantially altered by diagenesis^{23,25}.

Cibicides kullenbergi, *C. wuellerstorfi* and *Uvigerina* spp. ($\sim 0.5 \text{ mg}$ samples) from 14 late Miocene and 2 early Pliocene sediment samples were thoroughly cleaned using methods developed for Pleistocene samples¹⁷. Cadmium and manganese were quantified by graphite furnace atomic absorption spectroscopy, and calcium, by flame atomic absorption spectroscopy. Sample identifications, depths below the seafloor, ages and Cd/Ca and Mn/Ca results are presented in Table 1. To test

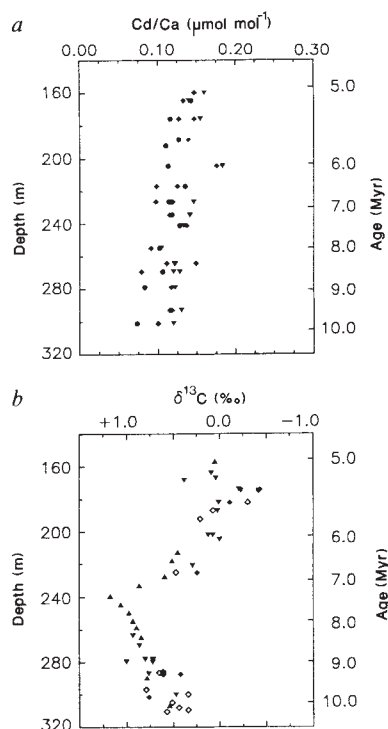


Fig. 1 *a*, Cd/Ca in benthic foraminifera from Site 289 plotted as a function of depth in sediment column. Plot symbols are: ▼, *C. wuellerstorfi*; ◆, *C. kullenbergi*; ●, *Uvigerina* spp. A small number of analyses known to be erroneous (results of small samples which do not agree with replicates, errors in sample handling, or samples with high Mn/Ca) are excluded from this plot, but included in Table 1 with a notation. *b*, $\delta^{13}\text{C}$ in benthic foraminifera from Site 289 plotted as a function of depth in sediment column. Plot symbols as in *a* with additions ▲, mixed *C. wuellerstorfi* and *C. kullenbergi* and ◇, *Cibicidoides* spp. Data from Savin *et al.*⁸.

whether Cd/Ca results could have been compromised by manganese-rich carbonate overgrowths²⁶, a larger sample of *Uvigerina* spp. (from 234.4 m) was sequentially dissolved with three aliquots of dilute acid (Table 1). An approximately constant Cd/Ca ratio with increasing dissolution, with a weighted average of 0.13 ± 0.01 , is consistent with results of total sample dissolutions from the same depth. The slight increase of Cd/Ca with cumulative fraction dissolved may be an artefact of adsorption of cadmium dissolved in early fractions onto surfaces of remaining solids, rather than a true increase in lattice-bound cadmium. The Mn/Ca ratio decreases by approximately a factor of two with increasing dissolution, probably from dissolution in the early fraction of minor amounts of manganese-rich carbonate overgrowths. The lack of higher Cd/Ca at elevated Mn/Ca during partial dissolution indicates that these moderate levels of Mn/Ca do not compromise Cd/Ca results.

The Cd/Ca ratios for these late Miocene benthic foraminifera are intermediate between those for Atlantic and Pacific samples observed back to interglacial Stage 7 (~215,000 yr) (ref. 17). Although several sites covering the most important ocean basins are needed to estimate global inventories, the present observation suggests that mean oceanic Cd during the late Miocene may have been roughly comparable to that of the present. This result attests to a remarkable stability of oceanic nutrient chemistry over long time intervals and despite major changes in global climate.

These late Miocene samples of *C. kullenbergi*, *C. wuellerstorfi* and *Uvigerina* spp. do not display significant differential fractionation for Cd/Ca (a result previously documented for Pleistocene samples¹⁷). *Cibicidoides* spp. (including *C. wuellerstorfi*)

Table 1 Site 289 (Leg 30) Cd/Ca and Mn/Ca

Core/sec/int	Depth (m)	Epoch	Age* (Myr)	Taxon†	Cd/Ca (×10 ⁻⁶)	Mn/Ca (×10 ⁻⁶)	Code‡
17/6/40-42	159.90	E Plio	5.09	uvi kul wue	0.068 0.147 0.160	59 — —	a
18/3/25-27	164.75	E Plio	5.19	uvi kul wue	0.143 0.133 0.140	28 66 51	a
19/4/40-42	175.90	L Mioc	5.41	uvi kul kul wue	0.116 0.127 0.147 0.155	19 36 41 64	
20/6/40-42	188.40	L Mioc	5.63	uvi kul wue	0.127 0.139 0.221	42 56 104	b
21/3/40-42	191.90	L Mioc	5.69	uvi kul	0.110 0.295	37 56	c
22/4/40-42	204.40	L Mioc	6.04	uvi kue wue	0.113 0.176 0.183	20 54 62	
23/6/40-42	216.90	L Mioc	6.52	uvi kul kul	0.135 0.125 0.098	46 43 19	
24/6/40-42	226.40	L Mioc	6.87	uvi uvi kul kul wue	0.117 0.114 0.097 0.119 0.146	37 48 42 55 66	a
25/5/40-42	234.40	L Mioc	7.18	uvi kul wue	0.118 0.115 0.141	47 34 53	
		Partial dissolution		Fraction dissolved:			
			48%	uvi	0.122	57	
			34%		0.130	41	
			18%		0.146	32	
26/3/40-42	240.90	L Mioc	7.44	uvi kul wue	0.128 0.137 0.133	42 55 49	
27/6/40-42	254.90	L Mioc	8.00	uvi kul wue	0.102 0.091 0.104	29 33 24	
28/6/40-42	264.40	L Mioc	8.38	uvi kul kul wue wue	0.055 0.149 0.111 0.121 0.122	21 36 43 36 37	a
29/3/40-42	269.40	L Mioc	8.58	uvi kul wue wue wue	0.106 0.079 0.120 0.128 0.083	16 28 32 33 21	
30/3/40-42	278.90	L Mioc	8.96	uvi kul wue wue	0.083 0.117 0.122 0.115	21 33 30 34	
31/6/40-42	292.90	L Mioc	9.52	kul kul wue	0.115 0.118 0.130	34 59 47	a
32/5/40-42	300.90	L Mioc	9.84	uvi kul wue	0.073 0.100 0.120	58 58 56	

* The depths of transitions between foraminiferal zones and of first and last appearance datum levels for Site 289 (from 19, 20) were assigned absolute ages based on the timescale of Berggren *et al.*, the correlations of foraminiferal biostratigraphic zones with this timescale, and the absolute ages assigned to first and last appearance datum levels³³. Ages for samples were interpolated between these depths and ages.

† Taxon; uvi, *Uvigerina* spp.; kul, *C. kullenbergi*; wue, *C. wuellerstorfi*.

‡ Codes are as follows: a, small sample, excluded if it does not agree with replicates; b, high Mn/Ca, excluded; c, error in sample handling, excluded.

and *Uvigerina* have significant offsets in carbon and oxygen isotope ratios, with *Uvigerina* isotopically lighter in carbon and heavier in oxygen^{8,27-29}. Calcification at different depths in the sediment for different taxa, and therefore from interstitial water of differing chemical composition due to organic matter regeneration, has been suggested to account for the carbon isotope offsets³⁰. Interstitial waters are enriched in Cd, apparently in proportion to nutrients from the decomposition of organic matter^{31,32}. Although more interstitial water measurements are

needed to determine the exact relationship between Cd increases and isotopically lighter carbon, the offsets in carbon isotopes for benthic foraminifera probably are not due to an infaunal habit.

Cd/Ca and *Cibicidoides* spp. $\delta^{13}\text{C}$ data (from ref. 8) for Site 289 are plotted as a function of depth and age in Fig. 1a and b. The ~5 Myr Cd/Ca record presented here does not indicate any major changes, even during times when major carbon isotope events occur. Oceanic residence times for Cd and C are probably similar (~100 kyr); responses to mass balance perturbations should occur at similar rates.

Loutit *et al.*¹¹ identified seven events (their MCI-MC7) in this portion of the late Miocene carbon isotopic record of Site 289, with the late Miocene carbon shift (MC2) placed between 240–220 m. Palaeomagnetic dating on other sites (with absolute age–biostratigraphy correlations revised to the Berggren *et al.*³³ timescale when necessary) has placed the late Miocene carbon shift in the upper part of chronozone 6, with assigned ages of ~6.2 Myr^{3–5,16} and 6.58–5.9 Myr (presented with arguments that age assignments or sampling at other sites have missed the full range of the shift¹⁵). The abrupt late Miocene carbon shift probably reflects the superposition of a low-frequency cyclic $\delta^{13}\text{C}$ cycle with a period of ~2–3 Myr and a long-term decrease in $\delta^{13}\text{C}$ (ref. 16). The more gradual apparent change at Site 289 may be a result of wider sample spacing. In addition, Site 289 was not palaeomagnetically dated, so absolute ages assigned from biostratigraphic correlations are less precise.

Cd/Ca and $\delta^{13}\text{C}$ records (Fig. 1) were averaged over three intervals—before, during and post-shift (Table 2). Average Cd/Ca appears to be slightly higher in the youngest post-shift section with the lightest carbon, but standard deviations of the mean for Cd/Ca overlap for all intervals. Linear regressions of Cd/Ca against depth (for the three taxa individually and pooled) yield negative slopes distinguishable from zero at 99.95% confidence limits for the pooled regression and at 99.5% for the individual regressions; these are equivalent to a decrease in Cd/Ca of $\sim 0.02 \pm 0.01$ over 60 m.

Three causes, acting individually or together, have been proposed to account for the late Miocene negative shift in carbon isotopes and cannot be distinguished solely on the basis of benthic carbon isotope data. The slope of the relationship between $\delta^{13}\text{C}$ and Cd/Ca in the ocean depends on $\Delta\delta^{13}\text{C}/\Delta\text{P}$ —the change in $\delta^{13}\text{C}$ per change in phosphorus—and on oceanic P/Cd; the intercept depends on oceanic mean $\delta^{13}\text{C}$ and P and $\Delta\delta^{13}\text{C}/\Delta\text{P}$ ¹⁸. The Cd/Ca data from benthic foraminifera can be used together with carbon isotope data to evaluate the role of each cause by assuming: (1) that the Redfield type regeneration ratio $\Delta\delta^{13}\text{C}/\Delta\text{P}$ was similar in the Miocene to that now; (2) that the oceanic distributions of Cd, P and $\delta^{13}\text{C}$ were linearly related; and (3) that P and Cd are depleted to near-zero levels in subtropical surface waters¹⁸.

One explanation ascribes the event to deep-ocean circulation changes, that is, initiation of or increases in formation of deep water similar to present-day North Atlantic Deep Water^{2,6,9,14}. This change would increase the age contrast between Atlantic and Pacific deep waters, thus enhancing nutrient and carbon isotope differences between the ocean basins. If the Miocene relationship between $\delta^{13}\text{C}$ and Cd was similar to those for the present and glacial stage two¹⁸, a decrease of 0.9% in $\delta^{13}\text{C}$ should result in an increase in Cd/Ca of 0.095. The observed change in mean Cd/Ca is only 0.03 ± 0.01 , less than a third of the expected increase. The Cd/Ca record therefore shows that deep-water circulation is not the major cause of the shift. This conclusion is supported by observations of a simultaneous negative carbon shift in benthic and planktonic foraminifera from both oceans^{6,8,9,14–16}.

A second explanation^{2,6,14} proposed to account for some or all of the shift to more negative benthic $\delta^{13}\text{C}$ is an increase in surface productivity stimulated by higher deep-water nutrient contents, which would make deep water $\delta^{13}\text{C}$ more negative

Table 2 Site 289 averaged Cd/Ca and carbon isotope values

Depth zone (m)	Cd/Ca (n)*	$\delta^{13}\text{C}$ (n)
170–210	0.139 ± 0.026 (10)	-0.12 ± 0.22 (14)
210–241	0.124 ± 0.015 (13)	0.57 ± 0.31 (8)
242–301	0.110 ± 0.019 (19)	0.76 ± 0.18 (22)
302–311		0.48 ± 0.09 (5)
All late Miocene	0.121 ± 0.022 (42)	0.45 ± 0.43 (49)

* n is number of values averaged.

relative to the surface ocean. This change might also enhance the 'vital effect' offsets for benthic foraminifera. But Cd/Ca data do not indicate taxa-specific offsets at any time, and it is unlikely that increased deep-water nutrients would occur with no change in deep Cd. Furthermore, parallel changes in benthic and planktonic carbon isotope records are difficult to reconcile with this cause^{6,16}. Finally, it must be assumed that none of the extra organic carbon produced is preserved in the sediments, or that the extra organic matter production is matched by a proportional increase in carbonate sedimentation; otherwise, the average $\delta^{13}\text{C}$ of the ocean would increase.

A third potential mechanism for the carbon shift is an increased proportion of ^{13}C -depleted organic carbon (relative to inorganic carbonate) during global continental weathering and/or decreased deposition of organic carbon during sedimentation^{2,6,9,10,13,14,16}. Shifts in organic carbon deposition and weathering are often linked to marine transgressions and regressions^{10,12,13}. The relationship between mean $\delta^{13}\text{C}$ and Cd/Ca places constraints on the characteristics of what was added to (or not removed from) the ocean. A shift in $\delta^{13}\text{C}$ without an accompanying Cd/Ca shift (assuming that the $\delta^{13}\text{C}$ change per change in phosphorus for the ocean remained about the same) could have resulted from a higher weathering rate of organic carbon, but this organic material must have been Cd-deficient relative to Cd/P and Cd/C ratios in typical marine organic matter. This material may have also been deficient in P relative to C compared to typical marine organic matter, but parallel shifts in planktonic and benthic $\delta^{13}\text{C}$ require that the oceanic P/C ratio remained approximately constant. Effecting a permanent shift in $\delta^{13}\text{C}$ from changes in organic carbon weathering or deposition rates requires coupled variations in other deposition and weathering rates (for instance, phosphorus, calcium carbonate). For example, increasing the weathering rate of organic carbon may also increase the P flux to the ocean. Without an increase in another P sink, potentially unrealistic increases in oceanic P inventories could occur.

This preliminary study adds significant constraints to models for the late Miocene carbon shift, one event in the temporal carbon isotope record. Late Miocene Cd/Ca data from other sites (and for middle and early Miocene, for example) are necessary to estimate changes in global oceanic mean Cd over time and to diagnose their relationship to $\delta^{13}\text{C}$ changes. Additionally, planktonic foraminiferal Cd/Ca data could confirm the assumption of near-zero nutrient levels in surface waters, and offer additional proof that the record is not diagenetically altered. As global mean Cd/Ca does not appear to change drastically, Cd may be the best tracer of changes in deep-water circulation patterns for the Cenozoic. High-resolution Cd/Ca work, similar to that done in the Pleistocene, would allow the determination of circulation changes during particularly interesting periods.

The Deep Sea Drilling Project samples were supplied through the assistance of the NSF. This work was supported by the NSF.

Received 5 June, accepted 14 September 1987.

1. Keigwin, L. D. Jr *Earth planet Sci. Lett.* **45**, 361-382 (1979).
2. Bender, M. L. & Keigwin, L. D. Jr *Earth planet Sci. Lett.* **45**, 383-393 (1979).
3. Loutit, T. S. & Kennett, J. P. *Science* **204**, 1196-1199 (1979).
4. Keigwin, L. D. Jr & Shackleton, N. J. *Nature* **284**, 613-614 (1980).
5. Haq, B. U. et al. *Geology* **8**, 427-431 (1980).
6. Vincent, E., Killingley, J. S. & Berger, W. H. *Mar. Micropaleont.* **5**, 185-203 (1980).
7. Woodruff, F. M., Savin, S. M. & Douglas, R. G. *Science* **212**, 665-668 (1981).
8. Savin, S. M. et al. *Mar. Micropaleont.* **6**, 423-450, (1981).
9. Bender, M. L. & Graham, D. W. *Mar. Micropaleont.* **6**, 451-464 (1981).
10. Loutit, T. S. & Keigwin, L. D. Jr *Nature* **300**, 163-166 (1982).
11. Loutit, T. S., Pisias, N. G. & Kennett, J. P. *Earth planet. Sci. Lett.* **66**, 48-62 (1983).
12. Loutit, T. S., Kennett, J. P. & Savin, S. M. *Mar. Micropaleont.* **8**, 215-233 (1983/84).
13. Woodruff, F. M. & Savin, S. M. *Geology* **13**, 119-122 (1985).
14. Vincent, E., Killingley, J. S. & Berger, W. H. *Mem. geol. Soc. Am.* **163**, 103-130 (1985).
15. Hodel, D. A. & Kennett, J. P. *Palaeoceanography* **1**, 285-311 (1986).
16. Keigwin, L. D., Aubry, M.-P. & Kent, D. V. *Init. Rep. DSDP Leg 94*, 935-963 (1987).
17. Boyle, E. A. & Keigwin, L. D. *Earth planet. Sci. Lett.* **76**, 135-150 (1985/86).
18. Boyle, E. A. *Geochim. cosmochim. Acta* **50**, 265-276 (1986).
19. Srinivasan, M. S. & Kennett, J. P. *Mar. Micropaleont.* **6**, 499-533 (1981).
20. Srinivasan, M. S. & Kennett, J. P. *Spec. Publ. Soc. econ. Paleont. Miner* **32**, 395-432 (1981b).
21. Woodruff, F. M. & Douglas, R. G. *Mar. Micropaleont.* **6**, 617-632 (1981).
22. Shackleton, N. J. *Prog. Oceanogr.* **11**, 199-218 (1982).
23. Elderfield, H. et al. *Geochim. cosmochim. Acta* **46**, 2259-2268 (1982).
24. Baker, P. A., Gieskes, J. M. & Elderfield, H. J. *sed. Petrol.* **52**, 71-82 (1982).
25. Delaney, M. L. & Boyle, E. A. *Earth planet. Sci. Lett.* **80**, 91-105 (1986).
26. Boyle, E. A. *Geochim. cosmochim. Acta* **47**, 1815-1819 (1983).
27. Woodruff, F. M., Savin, S. M. & Douglas, R. G. *Mar. Micropaleont.* **5**, 3-11 (1980).
28. Graham, D. W., Corliss, B. H., Bender, M. L. & Keigwin, L. D. *Mar. Micropaleont.* **6**, 483-497 (1981).
29. Duplessy, J.-C. *Quat. Res.* **21**, 225-243 (1984).
30. McCorkle, D. C., Emerson, S. R. & Quay, P. D. *Earth planet. Sci. Lett.* **74**, 13-26 (1985).
31. Klinkhammer, G., Heggie, D. T. & Graham, D. W. *Earth planet. Sci. Lett.* **61**, 211-219 (1982).
32. Heggie, D., Kahn, D. & Fischer, K. *Earth planet. Sci. Lett.* **80**, 106-116 (1986).
33. Berggren, W. A., Kent, D. V. & van Couvering, J. A. in *The Chronology of the Geological Record* (ed. Snelling, N. J.) *Mem. geol. Soc. Lond.* **10**, 211-260 (1985).

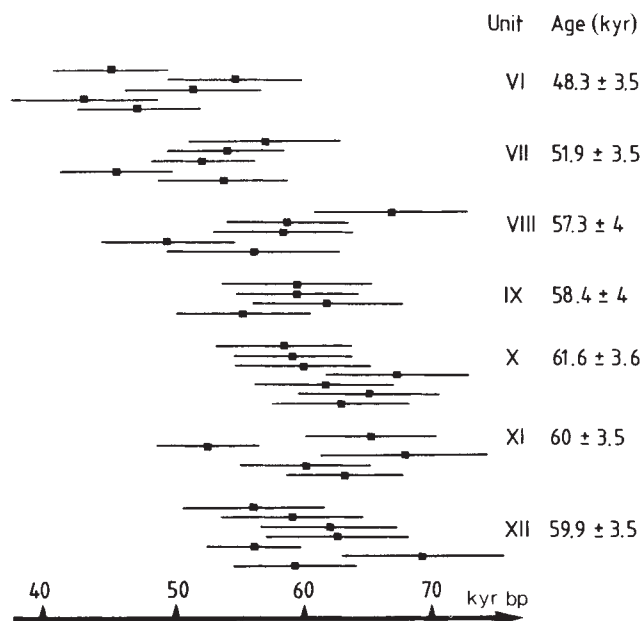


Fig. 1 Horizontal bars representing the thermoluminescence-dating ages of burnt flints from Kebara cave as a function of depth of the unit in which the flints were found. The weighted average age calculated for each unit is given at the right. The Mousterian sequence is about 4-m thick.

Thermoluminescence dates for the Neanderthal burial site at Kebara in Israel

H. Valladas*, J. L. Joron†, G. Valladas*, B. Arensburg‡, O. Bar-Yosef§, A. Belfer-Cohen§, P. Goldberg§, H. Laville||, L. Meignen||, Y. Rak‡, E. Tchernov #, A. M. Tillier** & B. Vandermeersch**

* Centre des Faibles Radioactivités, Laboratoire Mixte CNRS-CEA, Avenue de la Terrasse, 91190 Gif sur Yvette, France

† Groupe des Sciences de la Terre, Laboratoire Pierre Süe, CEN, Saclay, 91191 Gif sur Yvette, France

‡ Department of Anatomy and Anthropology, Sackler School of Medicine, Tel-Aviv University, Ramat Aviv 69978, Israel

§ Institute of Archaeology, Mount Scopus, Hebrew University, Jerusalem 91905, Israel

|| Institut du Quaternaire, Université de Bordeaux 1, 33405 Talence, France

¶ Centre de Recherches Archeologiques, Sophia Antipolis, 06565 Valbonne, France

Department of Zoology, Hebrew University, Jerusalem 91904, Israel

** Laboratoire d'Anthropologie, Université de Bordeaux 1, UA376 du CNRS, 33405 Talence, France

The origins of modern man are a subject of controversy among palaeoanthropologists concerned with human evolution¹⁻³. Particularly heavily debated is the dating of hominid remains uncovered in southwestern Asia, because middle palaeolithic sites have provided skeletal remains classified as representing Neanderthals (*Homo sapiens neanderthalensis* at Tabun, Amud, Kebara and Shanidar caves) and proto-Cro-Magnons (*Homo sapiens sapiens* at Skhul and Qafzeh caves). This situation differs considerably from that of Western Europe, where only Neanderthal remains are known from archaeological deposits of this period, or that of the African continent, where no Neanderthal remains have so far been found. Two opposing hypotheses have been offered to explain the relations between Neanderthals and the earliest modern *Homo sapiens*: first that modern *Homo sapiens* appeared very early in

the Mediterranean Levant and coexisted with a population of Neanderthals who had arrived at a later date; and second that modern humans developed from the local Neanderthal population in southwestern Asia. Recent excavations at the Kebara cave yielded Neanderthal burial in a well-documented stratigraphic and cultural Mousterian sequence^{4,5}. We now report that thermoluminescence dates from 38 specimens of burnt flint recovered from 4 m of Kebara deposits range from about 60,000 to 48,000 years before present (BP), indicating that Neanderthals were present in the Levant in the latter part of the middle Palaeolithic.

Current controversies about the geochronology of the middle Palaeolithic in Southwestern Asia include disagreements about the chronological interpretations of the Lebanese shorelines, the geological and cultural sequence of the Tabun cave, and the micro-mammalian bio-stratigraphical comparisons from Mousterian sites in the Levant^{5-11,28}. This situation motivated a new series of excavations at Kebara cave in 1982. Among the immediate aims of the fieldwork was to obtain thermoluminescence dates¹²⁻¹⁷ for the burnt flints while the excavation was in progress.

The Kebara cave (18 × 25 m) is on the steep western slope of Mount Carmel facing the Mediterranean Sea, about 30 km south of Haifa and about 15 km south of the Tabun cave; the Qafzeh cave is about 35 km to the east, in southern Galilee. The stratigraphic sequence at Kebara as exposed today contains 14 geoarchaeological units. The lowest, unit XIV, which in places rests upon the bedrock, is essentially sterile, consisting of sandy silts. Units XIII to VII appear to have accumulated quite rapidly with intermittent Mousterian occupations indicated by numerous superimposed hearths, *in situ* production of lithic artefacts, the occurrence of animal bones and the introduction of wind-blown silts and sands. In 1983, unit XII yielded a well-organized, almost complete burial of an adult Neanderthal^{4,5}. Units V and VI are Mousterian, but owing to the disturbing effects of erosion and animal burrowing contain a few intrusive upper Palaeolithic elements. Unit IV contains as yet unidentified industry, and units III to I an Upper Palaeolithic blade industry. Preliminary analysis of the lithic assemblages of units VII to XII indicates a great similarity between those of Kebara and those ascribed to the late Levantine

Table 1 Weighted mean ages of units in the Kebara sequence

Units	Mean age (BP)
VI	48,300 ± 3,500
VII	51,900 ± 3,500
VIII	57,300 ± 4,000
IX	58,400 ± 4,000
X	61,600 ± 3,600
XI	60,000 ± 3,500
XII	59,900 ± 3,500

The β -equivalent dose rate per unit of α -flux was measured as in ref. 22 and the β -equivalent annual dose of flints was calculated using specific fluxes deduced from ref. 23. The β annual dose in flints was calculated using the specific values given in ref. 24. The internal γ -dose is <0.01 mGy yr⁻¹. The external dose includes a cosmic contribution of $\sim 2 \times 10^{-2}$ mGy yr⁻¹, and was calculated, using the data given in refs 25 and 26. The thickness of overburden is about 20 m. (The detailed experimental data are available from H. Valladas.) The errors (statistical and systematic) at 68% confidence level have been assessed as in ref. 27. A systematic error of $\pm 7.5\%$ has been assumed for the external dose to take into account the possible variation of water content of the archaeological sediments (M. J. Aitken, personal communication).

Mousterian, defined on the basis of Tabun, layer B (refs 8 and 9).

The burnt flints were collected within and near the hearths of the trench in the central area of the excavation (units XII to X) and in the great western cut (units IX to VI) at the mouth of the cave. So far, the deepest Mousterian levels (units XIV and XIII) and those assigned to the Upper Palaeolithic have not yielded any burnt flints. Generally, five flints per unit were selected except for units XII and X, from each of which seven specimens were taken for dating. The technique used for the measurements of the palaeodose and the annual dose of radiation received by the flints have been fully described elsewhere¹⁵⁻¹⁷. The internal radiation dose of each flint was estimated from its uranium, thorium, and potassium concentrations which were determined by neutron activation analysis at the Pierre Sue Laboratory at Saclay¹⁸. The environmental dose rates of γ -rays and cosmic rays were measured using 45 dosimeters (CaSO₄/Dy) which had been buried for one year within each unit (~ 50 cm thick) near the spots where the flints were recovered. The radiation dose decreased with increasing depth from about 0.9 mGy yr⁻¹ in unit VI to 0.5 mGy yr⁻¹ in units XI and XII. The external dose of each flint was computed from the average recorded by the surrounding dosimeters; within each unit, the environmental dose did not vary by more than $\pm 10\%$. For the different specimens, the external dose represented from 40% to 70% of the total annual dose, which ranged from 0.70 to 1.40 mGy yr⁻¹.

In Fig. 1, the age of the 38 flints (with the statistical errors shown as horizontal bars) are plotted against the depth of recovery. The average ages estimated for the different units (Table 1) range from about 60,000 (units XII to X) to 48,000 (unit VI) yr BP and suggest that they were deposited during isotopic stage 3 (ref. 19). The ages calculated for units XII to X show no systematic variation with depth, suggesting a period of frequent occupation of the cave at some time $\sim 60,000$ BP. Although the average age of unit VII seems to be less than that of units VIII to XII, in view of the experimental error observed, it cannot be positively concluded that the occupation of the cave was interrupted for an extended period between levels VIII and VII. Consequently, the thermoluminescence data support the suggestion, made on the basis of stratigraphic evidence, that units XII to VII accumulated rapidly. The age proposed for unit VI must be treated with caution, because the flints recovered from it might have been brought from elsewhere by past erosion of which the great western cut shows ample evidence.

The thermoluminescence data presented here lead to two important conclusions. First, they suggest dates in the order of

50 to 60 kyr BP for the Mousterian industry of Tabun B-type, which are earlier than previously supposed⁸; we believe that this conclusion would hold true even if the Tabun C-type and B-type lithic industries are grouped together. Second, the date of 60,000 yrs BP for Neanderthal remains from Kebara fits the hypothesis that the Neanderthals were latecomers to South-western Asia^{20,21}. If the Neanderthal woman found in layer C of Tabun is of the same age or slightly older than the Kebara skeleton, as indicated by the stratigraphic position of the lithic industries, it adds further support to the idea that the Neanderthals made a relatively late arrival in the Levant. This would make the hypothesis of a close phylogenetic relationship between the Neanderthals and the proto-Cro-Magnons unlikely. However, this can only be resolved by the independent dating of the Qafzeh hominids, who, on bio-stratigraphic evidence, are older than the Levantine Neanderthals.

We thank the French Ministry of Foreign Affairs, the CNRS, the CEA and the L.S.B. Leakey Foundation for supporting the field and laboratory work.

Received 20 August; accepted 18 September 1987.

1. Vandermeersch, B. *Les hommes fossiles de Qafzeh, Israel* (CNRS, Paris, 1981).
2. Trinkaus, E. in *The Origins of Modern Humans* (eds Smith, F. H. & Spencer, F.) 251-293 (Lif, New York, 1984).
3. Wolpoff, M. H., Smith, H. F., Males, M., Radovic, J. & Rukavina, D. *Am. J. Phys. Anthropol.* **54**, 499-545 (1981).
4. Bar-Yosef, O. *et al. Curr. Anthropol.* **27**, 63-64 (1986).
5. Arensburg, B. *et al. C. r. heb. Seanc. Acad. Sci., Paris* **300**, 227-230 (1985).
6. Farrand, W. R. *J. Archaeol. Sci.* **6**, 369-392 (1979).
7. Salanville, P. in *Préhistoire du Levant* (eds Cauvin, J. & Salanville, P.) 21-32 (CNRS, Paris, 1981).
8. Jelinek, A. J. in *Préhistoire du Levant* (eds Cauvin, J. & Salanville, P.) 265-280 (CNRS, Paris, 1981).
9. Jelinek, A. J. *Science* **216**, 4553, 1369-1375 (1982).
10. Bar-Yosef, O. & Vandermeersch, B. in *Préhistoire du Levant* (eds Cauvin, J. & Salanville, P.) 281-286 (CNRS, Paris, 1981).
11. Tchernov, E. in *Quaternary Extinctions* (eds Martin, P. S. & Klein, R. G.) 528-552 (University of Arizona Press, Tucson, 1984).
12. Huxtable, J. & Jacobi, R. M. *Archaeometry* **24**, 164-169 (1982).
13. Bowman, S. G. E. & Sieveking, G. de G. *Proc. R. Soc. Lond. B* **205**, 353-361 (1982).
14. Aitken, M. J. *Thermoluminescence Dating* (Academic, New York, 1985).
15. Valladas, H. thesis, Univ. Paris VI (1985).
16. Valladas, H., Geneste, J. M., Joron, J. L. & Chadelle, J. P. *Nature* **322**, 452-454 (1986).
17. Valladas, H. & Valladas, G. *Archaeometry* **27**, 214-220 (1987).
18. Chayla, B., Jaffrezic, H. & Joron, J. L. *C. r. heb. Seanc. Acad. Sci., Paris* **277**, 273-275 (1973).
19. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39-55, (1973).
20. Vandermeersch, B. in *The Transition from the Lower to the Middle Palaeolithic and the Origin of Modern Man* (ed. Ronen, A.) 297-300 (BAR, Oxford, 1982).
21. Bar-Yosef, O. in *Corridors, Cul-de-Sac and Coalescence: the Bio-cultural Foundations of Modern People* (ed. Trinkaus, E.) (Cambridge University Press, in the press).
22. Valladas, H. & Valladas, G. *Proc. R. Soc. Lond. B* **171**, 171-178 (1982).
23. Aitken, M. J. & Bowman, S. G. E. *Archaeometry* **17**, 132-138 (1975).
24. Bell, W. T. *Archaeometry* **21**, 243-245 (1979).
25. Kvasnicka, J. *High Phys.* **36**, 521-524 (1979).
26. Prescott, J. R. & Stephan, L. G. *Proc. R. Soc. Lond. B* **171**, 17-25 (1982).
27. Aitken, M. J. *Archaeometry* **18**, 233-238 (1976).
28. Gvirtzman, G. Annual Meeting of the Israel Society for Quaternary Research, 14-20 (University of Jerusalem, 1983).

A novel mechanism of insect resistance engineered into tobacco

Vaughan A. Hilder*, Angharad M. R. Gatehouse*, Suzanne E. Sheerman†, Richard F. Barker† & Donald Boulter*

*Department of Botany, University of Durham, South Road, Durham DH1 3LE, UK

†Plant Breeding Institute, Trumpington, Cambridge CB2 2LQ, UK

A major goal of plant genetic engineering is the introduction of agronomically desirable phenotypic traits into crop plants in situations where conventional breeding methods have been unsuccessful. One such target is enhanced resistance to insect pests which, in view of the estimated production losses world-wide and the heavy costs of protective treatments, is very important. We report here that a gene encoding a cowpea trypsin inhibitor, which has been shown to give some measure of field resistance to insect pests¹,

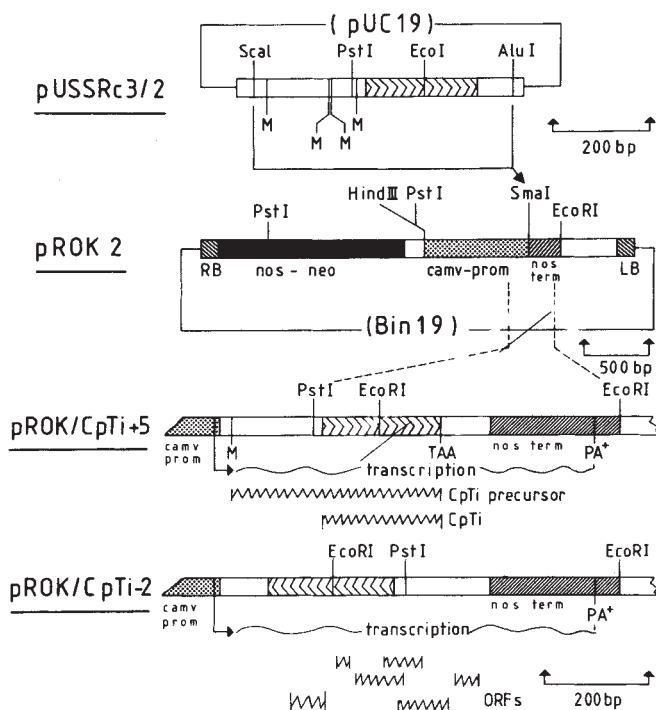


Fig. 1 Structure of the CpTI cDNA constructs. Details of the cloning and sequencing of pUSSRc3/2 will be presented elsewhere. Briefly, it was identified as a full length CpTI cDNA clone on the basis of cross-hybridization with a mixed synthetic oligonucleotide specific to a region of the major CpTI; with a previously isolated CpTI partial cDNA clone pAGCI; and by direct nucleotide sequencing. The coding sequence for the mature inhibitor [>>] and the upstream, in-phase ATG-methionine-initiator codons (M) are indicated. The 550-bp *AluI*-*ScaI* restriction fragment was blunt-end ligated into the *SmaI* site of the expression vector pRok2. Clones with a single cDNA insert in either orientation were identified by restriction mapping. In the correct orientation (pRok/CpTI+5) transcription from the CaMV promoter should be translated into CpTI precursor from the initiator codon ~20-nucleotides from the 5'-end of the messenger RNA. In the reverse orientation (pRok/CpTI-2) there are six short open reading frames.

confers, when transferred to tobacco, enhanced resistance to this species' own herbivorous insect pests.

Although the techniques for introducing foreign genes into many crop plants are now becoming routine, the identification of useful genes to be transferred remains a difficult problem. Examples which are of agricultural importance include herbicide resistance^{2,3} and virus resistance^{4,5}. Attempts to address this problem in relation to insect resistance have been made using the *Bacillus thuringiensis* endotoxin genes which have recently been introduced into tobacco⁶ and tomato⁷. The potential of these genes is limited since *B. thuringiensis* toxins are very specific. Also, no field resistance to insects of commercial formulations of the *B. thuringiensis* toxins have yet been reported, despite years of use⁸. In contrast, several general mechanisms for plant protection against insects exist naturally which confer field resistance to a wide range of pests, and our strategy has been to attempt to identify and exploit these. One such mechanism involves the trypsin inhibitors of cowpea (*Vigna unguiculata*) which have a number of attractive features for our present purposes. Cowpea trypsin inhibitors (CpTIs) are small polypeptides of around 80 amino acids belonging to the Bowman-Birk type of double-headed serine protease inhibitors and are the products of a small gene family⁹. As their metabolic target is the catalytic site of an enzyme, the ability of the insects to evolve a resistance mechanism based on mutation at this site should be minimal. The level of CpTI within the seeds of cowpeas correlated with field resistance to their major pest, the

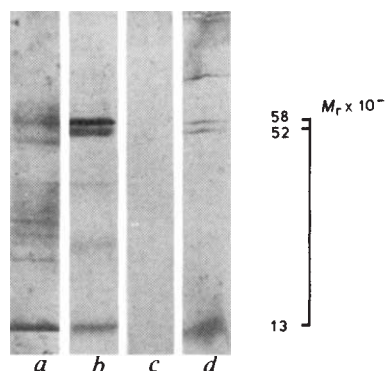


Fig. 2 Western blot analysis of transformed *N. tabacum* leaf tissue. *a*, Affinity-purified CpTI; *b*, seed extract of cowpea TVu 2027; *c*, -2/8 leaf tissue; *d*, +5/5 leaf tissue. Samples were run on SDS-polyacrylamide gel electrophoresis (17% gel slabs) under non-reducing conditions. Proteins were then transferred to nitrocellulose and reacted with polyclonal CpTI antibodies³⁵.

bruchid beetle, *Callosobruchus maculatus*¹. More importantly, feeding trials with purified CpTI, incorporated at physiological levels into artificial diets¹⁰, have shown these inhibitors to be anti-metabolic agents against a wide range of insects, including genera such as *Heliothis*, *Spodoptera*, *Diabrotica* and *Tribolium*, which cause losses of major economic importance. There is no suggestion that cowpea seeds are toxic to humans, and although normally cooked, they can be eaten raw¹¹. Furthermore, raw cowpea seed meal fed to rats did not affect their growth (D. Boulter, unpublished observations).

The CpTI gene which was used was derived from plasmid pUSSRc3/2, a member of a complementary DNA library prepared from cowpea cotyledon polyadenylated RNA (Fig. 1). The homology between this and most other reported plant proteinase inhibitor genes¹²⁻¹⁷ is not great. A 550-base-pair (bp) long *AluI*-*ScaI* restriction fragment containing the 240-bp coding sequence for the mature inhibitor, a long leader sequence with four in-phase methionine codons, and a 96-bp 3' non-translated sequence, was transferred to the *SmaI* site of *Agrobacterium tumefaciens* Ti plasmid binary vector^{18,19} pRok2 (provided by M. Bevan and T. Kavanagh). This is related to the pRok1 vectors⁴ from which it differs primarily in having a multipurpose cloning site between the strong, constitutive CaMV 35S gene promoter²⁰ and the nopaline synthase gene-transcription termination sequence²¹. Clones were identified which contained an insert in the correct orientation relative to the CaMV promoter (pRok/CpTI+5) to produce CpTI, and in the 'reverse' orientation (pRok/CpTI-2), which has six short open reading frames with no identifiable features. This 'reversed' construct was used to produce control transformants. These constructs were mobilized into *Agrobacterium tumefaciens*¹⁸ and used to transform leaf discs of *Nicotiana tabacum* (cv. Samsun NN). Transformants were selected by their resistance to kanamycin, and transformed plants were regenerated from shootlets by transfer to a root-inducing, kanamycin-containing agar medium²². Rooted plants were grown on in soil-based compost. The plants showed no phenotypic abnormalities.

We are concerned here with the phenotypic effects of expression of the CpTI polypeptide in tobacco plants on insect resistance. Details of the molecular basis of these phenotypes will be presented elsewhere. The plants we examined in depth contained 3-7 unrearranged copies of the construct, which were simply and stably inherited. Levels of CpTI production in the original transformants were measured by dot-immunobinding assays²³ using polyclonal antibodies raised in rabbits against total CpTI. The level of expression in young leaves from different individual PRok/CpTI+5 transformants ranged from below the limit of detection to ~1% of total soluble protein (Table 1). This range in levels of expression was expected using the

Table 1 CpTI content and bioassay of transformed plants

Original plants*	CpTI detected† (μg per mg soluble protein)	Trypsin inhibitor activity‡	% Insect survival
-2/2	<1		100
-2/4	<1		87.5
-2/7	<1		100
-2/8	<1	1.7	87.5
+5/1	1.1		87.5
+5/3	1.3		100
+5/4	<1		100
+5/5	9.6	9.9	37.5
+5/9	5.8		62.5
+5/11	3.8		62.5
+5/14	4.2		75.0
+5/19	2.5		62.5
+5/21	6.2	7.0	50
+5/23	<1		100

Replicate plants	% Leaf area eaten §	Surviving insect biomass per plant (mg)
-2/8($n=11$)	52.3 $\pm$ 7.5	236.4 $\pm$ 33.3
+5/5($n=12$)	21.1 $\pm$ 1.4	63.9 $\pm$ 9.6
-2/8($n=10$)	49.6 $\pm$ 4.5	193.5 $\pm$ 10.2
+5/21($n=10$)	30.1 $\pm$ 1.5	109.8 $\pm$ 5.9

Bioassays were carried out as described in Fig. 2. Means and standard errors are provided for the bioassays on sets of clonally replicated plants.

* Only those plants used in two feeding trials are recorded.

† CpTI was detected using a dot immunobinding assay on soluble proteins extracted from recently fully expanded leaves which were removed shortly before infestation. Standards were affinity purified CpTI in control protein extracts from untransformed plants.

‡ Trypsin inhibitor activity was measured in leaf soluble protein extracts, according to the method of Erlanger *et al.*²⁶, expressed as μg equivalent of affinity purified CpTI per mg leaf protein. The leaves for this assay were recently expanded ones from more mature plants (after bioassay).

§ Leaf area and damage determined by image analysis of the third mature leaf from the apex (usually the most damaged).

CaMV 35S gene promoter^{5,7,19,24,25}. Expression of CpTI in pRoK/CpTI-2 transformants could not be detected.

Western blotting of soluble leaf proteins from the highest expressing plant, +5/5, shows that a polypeptide is produced and processed in this plant which corresponds to one of the isoforms from cowpea seed (Fig. 2). A mature protein band of relative molecular mass 13,000 (M_r) and higher M_r bands representing either trypsin inhibitor precursors or aggregates, are present in both +5/5 and cowpea seed extracts. None of these bands are present in -2/8. The functional integrity of the CpTI produced in these plants was demonstrated by *in vitro* trypsin inhibitor activity assay²⁶ (Table 1).

The critical test on these plants was the bioassay for insecticidal activity in insect feeding trials. Young plants were infested with newly emerged larvae of the lepidopteran *Heliothis virescens* and sealed into individual plantaria within the growth cabinet. *H. virescens*, the tobacco budworm, was selected as it is classed as a serious economic pest, one of whose primary hosts is tobacco. After seven days, all larvae (both dead and surviving) were removed, their size recorded and the extent of leaf damage measured. Plants which were not seriously damaged were subjected to further rounds of infestation. Considerable variability within the CpTI+5 and CpTI-2 populations was observed, as expected when dealing with two inherently variable biological systems (that is, plants and insects) with a small population of insects on each plant and with the variable efficacy of transferred genes. However, insect survival on, and damage to, ~20% of the CpTI+5 transformants was clearly decreased compared to

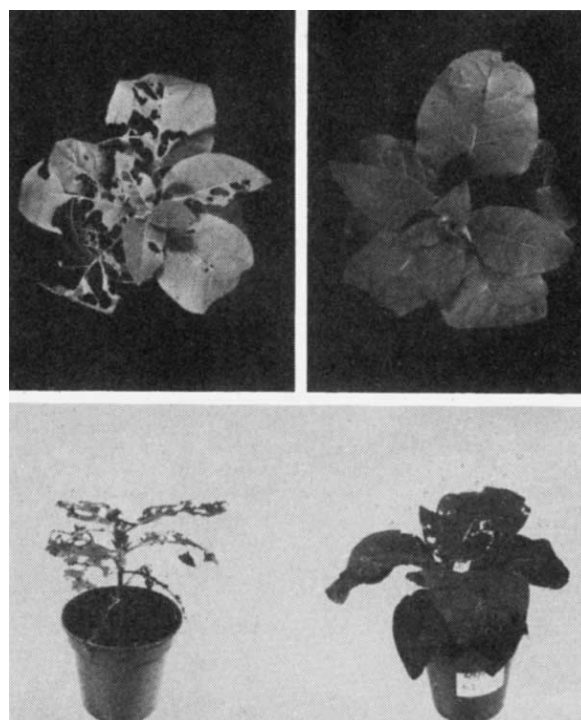


Fig. 3 Effect of *H. virescens* larvae on *N. tabacum* transformed with CpTI+5 and CpTI-2 constructs. The plant on the left is a clonal replicate of a CpTI-2 line; on the right, a CpTI-expressing, CpTI+5 line.

Methods. The pRoKcPcTI constructs were mobilized into *Agrobacterium tumefaciens* strain LBA 4404 in a triparental mating with *E. coli* HB101 harbouring pRK2013 (ref. 18). Transconjugants harbouring CpTI+5 and CpTI-2, as verified by restriction analysis, were used to transform leaf disks from axenically-grown *N. tabacum* and transformants were selected by inclusion of Kanamycin at 100 $\mu\text{g ml}^{-1}$ in the shoot-inducing and root-inducing media used to regenerate plants as described by Horsch *et al.*²². *H. virescens* was obtained from USDA, Brownsville, Texas, and maintained as described by Singh³⁶. Young, 10–12 cm tall plants were infested with eight newly emerged larvae (previously determined to be the optimum number for insect survival and development on control plants) and sealed into individual plantaria. Insects were left on the plants for seven days.

the controls in repeated trials (Fig. 3). These plants were the ones which expressed the highest levels of CpTI (Table 1).

Some of the transformants showing enhanced insect resistance and some of the controls were replicated as stem cuttings⁴ to provide sets of 12 genetically identical clonal plants on which statistically sound insect feeding trials could be run. These have provided convincing evidence that the CpTI-producing plants are much more resistant to insect attack (Table 1b). Control plants are devastated by this level of infestation; in trials which we have run beyond seven days, to 'termination', they were reduced to a stalk. On the CpTI-producing plants +5/5 and +5/21, although the larvae begin to feed and do some very limited damage to the leaves, they either die or fail to develop as they would on control plants, consistent with the predicted mechanism of CpTI toxicity¹⁰.

We have demonstrated that expression of a CpTI gene in tobacco leads to enhanced resistance to one of its major insect pests, at least in laboratory trials. Incorporation of this gene into other agronomically important crops, such as cotton, maize and rice, should soon be technologically feasible^{27–29}, although the ultimate effectiveness of this method can only be determined by testing for resistance in the field.

These experiments also shed light on the natural role of plant proteinase inhibitors. Proteinase inhibitors have been shown to

be present in the leaves of various species and, in some instances to be induced by insect attack^{30,31}. Although they are implicated in resistance to herbivorous insects³², direct evidence for the effectiveness of a given proteinase inhibitor against a specific insect has been lacking. Recent evidence suggests that they are one part of the complex interaction between plant nutritional value and the insects' digestive physiology³³. Our results indicate that a proteinase inhibitor can, by itself, significantly reduce insect attack.

Genetic engineering of crops offers opportunities to establish multi-line mixtures by creating new iso-genic lines containing different resistance genes, in what are otherwise identical genetic backgrounds; until now a strategy not pursued because it is extremely time-consuming to produce these by conventional methods. Such multi-lines, however, are likely to be effective since sowing intimate mixtures of contrasting varieties in the same experimental plots has led to slower build-up of disease³⁴. The use of such mixtures should form part of an integrated system of pest management, which would not necessarily exclude the judicious use of chemical protectants.

We thank Paulette Lopez for supplying cultures of *H. virescens*, Gillian Davison for technical assistance and John Gatehouse for discussions. Transfer of the CpTI coding fragment into *Agrobacterium* was performed by Dr. M. Bevan. Funding by Agricultural Genetics Company is gratefully acknowledged.

Received 11 August; accepted 1 October 1987.

1. Gatehouse, A. M. R. G., Gatehouse, J. A., Dobie, P., Kilminster, A. M. & Boulter, D. J. *Sci. Fd Agric.* **30**, 948-958 (1979).

2. Comai, L. *et al. Nature* **317**, 741-744 (1985).
3. Shad, D. *et al. Science* **233**, 478-481 (1986).
4. Baulcombe, D. C., Saunders, G. R., Bevan, M., Mayo, M. A. & Harrison, B. D. *Nature* **321**, 446-449 (1986).
5. Abel, P. P. *et al. Science* **232**, 738-743 (1986).
6. Vaecq, M. *et al. Nature* **328**, 33-37 (1987).
7. Fischhoff, D. A. *et al. Biotechnology* **5**, 807-813 (1987).
8. Shields, R. *Nature* **328**, 12-13 (1987).
9. Gatehouse, A. M. R. G., Gatehouse, J. A. & Boulter, D. *Phytochemistry* **19**, 751-756 (1980).
10. Gatehouse, A. M. R. G. & Boulter, D. *J. Sci. Fd Agric.* **34**, 345-350 (1983).
11. Peterson, V. in *The Natural Food Catalogue* **69**, (McDonald, London, 1984).
12. Hoffmann, L. M. Sengupta-Gopalan, C. & Paaren, H. E. *P. molec. biol.* **3**, 111-117 (1984).
13. Hammond, R. W., Foard, D. E. & Larkins, B. A. *J. biol. Chem.* **259**, 9883-9990 (1984).
14. Lee, J. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 7277-7281 (1986).
15. Cleveland, T. E., Thornburg, R. W. & Ryan, C. A. *P. molec. Biol.* **8**, 199-207 (1987).
16. Graham, J. S. *et al. J. biol. Chem.* **260**, 6561-6564 (1985).
17. Keil, M., Sanchez-Serrano, J., Schell, J. & Willmitzer, L. *Nucleic Acids Res.* **14**, 5641-5650 (1986).
18. Bevan, M. *Nucleic Acids Res.* **12**, 8711-8721 (1984).
19. Bevan, M., Mason, S. E. & Goelet, P. *EMBO J.* **4**, 1921-1926 (1985).
20. Guilley, H., Dudley, R. K., Jonard, G., Balazs, E. & Richards, K. E. *Cell* **30**, 763-773 (1982).
21. Bevan, M., Barnes, W. M. & Chilton, M. D. *Nucleic Acids Res.* **11**, 369-385 (1983).
22. Horsch, R. B. *et al. Science* **227**, 1229-1231 (1985).
23. Jahn, R., Schiebler, W. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1684-1687 (1984).
24. Sanders, P. R., Winter, J. A., Barnason, A. R., Rogers, S. G. & Fraley, R. T. *Nucleic Acids Res.* **15**, 1543-1558 (1987).
25. Jones, J. D. G., Dunsmuir, P. & Bedbrook, J. *EMBO J.* **4**, 2411-2418 (1985).
26. Erlanger, B. F., Kokowsky, N. & Cohen, W. *Archs Biochem. Biophys.* **95**, 271-278 (1961).
27. Umbeck, P., Johnson, G., Barton, K. & Swain, W. *Biotechnology* **5**, 263-266 (1987).
28. Pau, E., Mehra-Palta, A., Nagy, F. & Chua, N. *Biotechnology* **5**, 815-817 (1987).
29. Graves, A. C. F. Goldman, S. L. *P. molec. Biol.* **7**, 43-50 (1986).
30. Ryan, C. A. in *The Biochemistry of Plants: a Comprehensive Treatise*, vol. 6 (ed. Marcus, A.) 351-370 (Academic, New York 1985).
31. Brown, W. E., Takio, K., Titani, K. & Ryan, C. A. *Biochemistry* **24**, 2105-2108 (1985).
32. Ryan, C. A. in *Variable Plants and Herbivores in Natural and Managed Systems* (eds Denno, R. F. & McClure, M. S.) 43-60 (Academic, New York 1983).
33. Broadway, R. M., Duffrey, S. S., Pearce, G. & Ryan, C. A. *Ent. exp. appl.* **41**, 33-38 (1986).
34. Wolfe, M. S. & Barrett, J. A. *P. Disease* **64**: 148-155 (1980).
35. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
36. Singh, P. in *Artificial Diets for Insects, Mites and Spiders*. (IFI, Plenum, New York, 1982).

Postsynaptic fall in intracellular pH induced by GABA-activated bicarbonate conductance

K. Kaila & J. Voipio

Department of Zoology, Division of Physiology,
University of Helsinki, Arkadiankatu 7, SF-00100 Helsinki, Finland

Synaptic inhibition mediated by γ -aminobutyric acid (GABA) is known to involve opening of receptor-gated chloride channels¹⁻³. Recent evidence indicates that these channels also show a significant permeability to the physiologically important bicarbonate anion⁴. In all the excitable cells studied to date, the intracellular pH (pH_i) is higher than would be predicted from a passive distribution of H^+ ions⁵⁻⁹, and consequently there is an outwardly directed electrochemical driving force for HCO_3^- . In the presence of CO_2/HCO_3^- therefore, activation of GABA-gated channels could give rise to a significant efflux of bicarbonate, leading to a fall in postsynaptic pH_i . We have examined the influence of GABA on pH_i in crayfish skeletal muscle and we find that in the presence of CO_2 , GABA induces a dramatic fall in pH_i which is coupled to an alkalosis at the extracellular surface. This fall in pH_i and the extracellular alkalosis are attributable to a GABA-activated, picrotoxin-sensitive HCO_3^- -conductance. In view of the sensitivity of ion channels¹⁰ and intracellular ion concentrations⁵⁻⁹ to changes in pH_i , a GABA-induced postsynaptic acidosis could prove to be important in the modulation of inhibitory transmission.

Previous work on the dactyl opener carried out in the absence of HCO_3^- has shown that bath-applied GABA (10^{-5} - 10^{-3} M) produces a dose-dependent conductance increase due to opening of receptor-gated Cl^- channels^{3,11}. The effect of GABA shows little desensitization^{3,11}. In experiments of the type shown in Fig. 1, we simultaneously examined the effects of GABA on membrane potential, input resistance, intracellular Cl^- activity

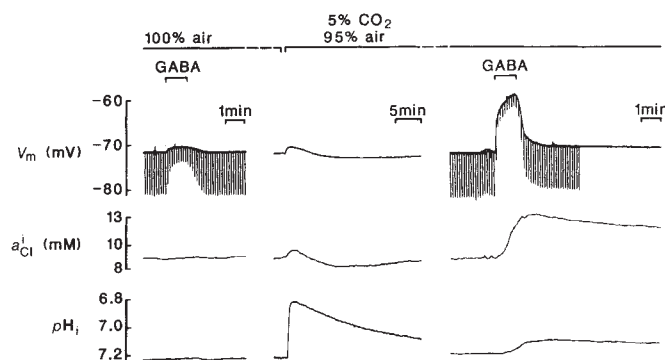


Fig. 1 Effects of bath-applied GABA (10^{-4} M) on membrane potential (V_m), on input resistance as measured by constant current pulses (100 nA, 0.3 s; two-electrode current clamp), on intracellular chloride activity (a_{Cl^-}) and on intracellular pH (pH_i). All four microelectrodes were inserted in the same fibre. The effects of GABA were first measured in a saline equilibrated with air (left). After recovery of a_{Cl^-} and pH_i from exposure to 5% CO_2 with 95% air (centre), GABA was reapplied (right). The CO_2 -free solution contained (mM): NaCl 200, KCl 5.4, $CaCl_2$ 7.0, $MgCl_2$ 2.6, HEPES 10 (pH 7.4); and the 5%- CO_2 solution contained (mM): NaCl 170, KCl 5.4, $CaCl_2$ 7.0, $MgCl_2$ 2.6, $NaHCO_3$ 30 (pH 7.4). The experimental temperature was 20 °C. a_{Cl^-} was estimated using a chloride activity coefficient of 0.73. Breaks in traces are 5 min and 35 min.

Methods. The experiments were done on the dactyl opener muscle of the first walking leg of the crayfish *Astacus astacus*. The individual muscle fibres ranged from 110 to 180 μm in diameter and 1.6-2.3 mm in length. The preparation was superfused²² at a rate of 8-10 bath volumes per second with crayfish saline equilibrated with 100% air, or with air containing either 5% or 1% CO_2 (the pH of all solutions was 7.4). Microelectrodes filled with 0.8 M K_2SO_4 + 8 mM KCl were used to record membrane potential (V_m) and to pass current. The intracellular Cl^- activity and intracellular pH were measured using liquid-sensor, ion-selective microelectrodes²³⁻²⁶.

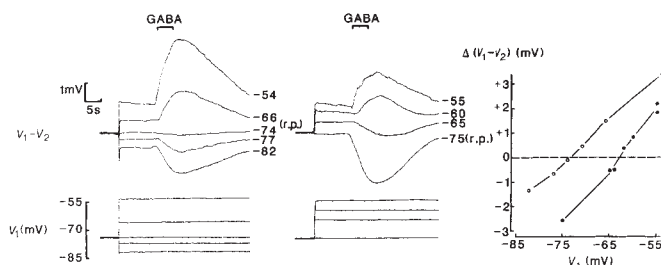


Fig. 2 Effect of $\text{CO}_2/\text{HCO}_3^-$ on E_{GABA} measured using a three-electrode voltage clamp. Voltage control was imposed on V_1 -electrode located in the centre of the fibre; V_2 -electrode was located at midway between centre and end. The current electrode was close to V_1 . Specimen recordings are shown in 100% air (left) and in the presence (middle) of 5% CO_2 + 95% air. The changes in $V_1 - V_2$, $\Delta(V_1 - V_2)$, following a 5 s pulse of GABA (5×10^{-4} M) in the absence (open circles) and presence of CO_2 (filled circles) have been plotted as a function of holding potential (right). Note that when GABA is applied in the presence of CO_2 , there is a large negative deflection in $V_1 - V_2$, indicative of an inward current at the level of the resting potential (r.p.). The steady-state pH_i (7.13–7.15) in the CO_2 -solution was not affected by the 5 s pulses of GABA (not shown). The intracellular HCO_3^- concentration was obtained from: $[\text{HCO}_3^-]_i = 10^{(\text{pH}_i - \text{pH}_o)} \times [\text{HCO}_3^-]_o$. The solutions were the same as those in Fig. 1 except for the addition of 3×10^{-3} M Mn^{2+} to block contractile activation by the depolarizing pulses. Control experiments showed that Mn^{2+} has no influence on the effects of GABA (10^{-5} – 10^{-3} M) on membrane potential and conductance.

(a_{Cl^-}) and pH_i . In the absence of HCO_3^- , the fall in membrane resistance caused by GABA (at 1 – 5×10^{-4} M) was paralleled by a very small depolarization (0.5–2.5 mV in 8 fibres) and no observable change in a_{Cl^-} . Effects of this kind are to be expected as the resting chloride conductance is known to be high in this preparation^{12,13}, and Cl^- ions are therefore at, or close to, equilibrium across the cell membrane¹⁴. In the experiment shown in Fig. 1, the chloride equilibrium potential (E_{Cl^-}) estimated from a_{Cl^-} was -72 mV in the absence of HCO_3^- .

Exposing the preparation to saline equilibrated with 5% CO_2 at constant external pH (7.4) resulted in a transient intracellular acidification with a small, transient decrease in a_{Cl^-} (probably due to an efflux of chloride on $\text{Cl}^-/\text{HCO}_3^-$ exchange⁸) coupled to the recovery of pH_i . If applied in the presence of HCO_3^- , GABA again produces a fall in membrane resistance which is now associated with a marked depolarization. This depolarization cannot be explained on the basis of a GABA-induced Cl^- -current because first, E_{Cl^-} before GABA application is virtually the same (-70 mV) as in the absence of HCO_3^- , and much more negative than the membrane potential (V_m) at peak depolarization; and second, the depolarization is paralleled by an increase in a_{Cl^-} , implying a net influx of chloride ions. These observations are best explained by assuming that GABA activates a conductance increase to both Cl^- and to some other ion, which has an equilibrium potential positive to V_m and E_{Cl^-} . On this scheme, the influx of Cl^- is due to passive redistribution of chloride caused by the depolarizing action of the unidentified ion, which our results (see below) would predict to be HCO_3^- . The depolarization caused by GABA (1 – 5×10^{-4} M) in 5% CO_2 saline was 13–17 mV (7 preparations) and was inhibited by 10^{-4} M picrotoxin (PTX), a blocker of the GABA-gated Cl^- channel¹⁵ (see Fig. 4). Figure 1 also shows that, in the presence of 5% CO_2 , GABA produces a decrease in pH_i , an effect consistent with a passive efflux of HCO_3^- ions down their electrochemical gradient. Note that at a constant partial pressure of CO_2 , the bicarbonate equilibrium potential $E_{\text{HCO}_3^-} = (\text{pH}_i - \text{pH}_o) \times 58$ mV which, at a pH_i 7.2 (Fig. 1), is -12 mV.

The effect of CO_2 saline on E_{GABA} was studied by means of a three-microelectrode voltage clamp^{16,17} while simultaneously

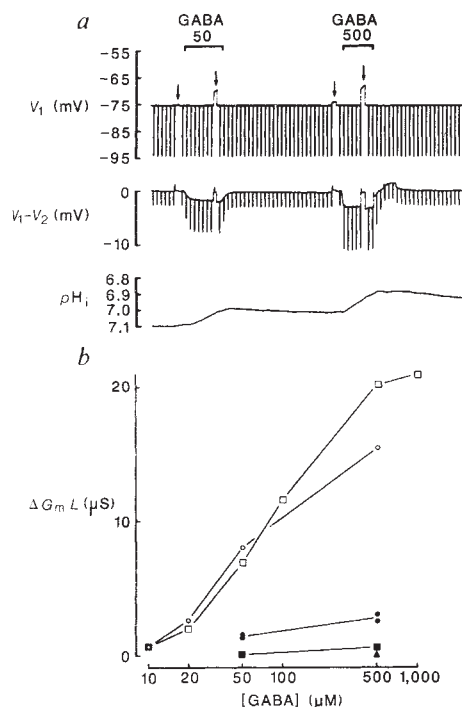


Fig. 3 Effects of GABA on membrane conductance, as measured using a three-microelectrode voltage clamp¹⁷, and on pH_i . All data shown are from a single, continuous experiment. *a*, Specimen recording of the effects of GABA (5×10^{-5} M and 5×10^{-4} M) in solution equilibrated with 5% CO_2 + 95% air is shown to illustrate method of measurement. Electrode placement as in Fig. 2, with the pH-selective electrode referred to and located close to V_1 -electrode. In all experiments of this type, $L/\lambda < 1.8$ and $(V_1 - V_2)/V_2 < 1.0$, where L is the half-length of the fibre and λ the space constant. V_1 was clamped close to the resting level and hyperpolarizing pulses (19 mV, 0.3 s) were applied at 0.08 Hz. Shortly before and during GABA application, the clamp was switched off for 8–16 s (arrows). Note decrease in pH_i and negative deflection in the base-line level of $V_1 - V_2$, indicative of an inward shift in holding current (and depolarization when off-clamp) following application of GABA. *b*, The increase in half-fibre conductance ($\Delta G_m L$) measured in Cl^- -containing (open symbols) and Cl^- -free solutions (filled symbols), equilibrated either with 100% air (squares) or 5% CO_2 + 95% air (circles), is plotted as a function of GABA concentration. PTX was applied at 10^{-4} M in the Cl^- -free, HCO_3^- -containing solution (triangle). All solutions contained 3×10^{-3} M Mn^{2+} to prevent contractions induced by withdrawal of chloride. The chloride-free solution was made by isomolar substitution of Cl^- with methanesulphonate.

measuring pH_i to allow estimation of intracellular bicarbonate concentration, $[\text{HCO}_3^-]_i$. GABA (5×10^{-4} M) was applied in pulses of five seconds and the membrane potential was clamped 8–11 seconds before GABA application. The measurements were repeated at intervals of 10 minutes, during which time the fibre was not clamped. Figure 2 shows that in a HCO_3^- -free medium, E_{GABA} is 1.6 mV positive with respect to the resting membrane potential (RP), and that in the presence of bicarbonate (50 minutes after start of the CO_2 exposure to allow complete recovery of pH_i), $E_{\text{GABA}} - \text{RP}$ is $+12$ mV. The respective values for three other fibres ranged from 0.5–2 mV and from 11–14 mV. The values of E_{GABA} and pH_i obtained in the CO_2 -solution yielded estimates for the permeability ratio $P_{\text{Cl}^-}:P_{\text{HCO}_3^-}$ from 1:0.46 to 1:0.53, as calculated assuming constant field conditions^{18,19}, that Cl^- is at equilibrium with resting V_m (see above), and that the GABA-induced current is carried only by Cl^- and HCO_3^- .

The ionic basis of GABA action was examined further (Fig. 3) by studying the effect of GABA on membrane conductance¹⁷ in the standard CO_2 -free and CO_2 -containing solutions, in a

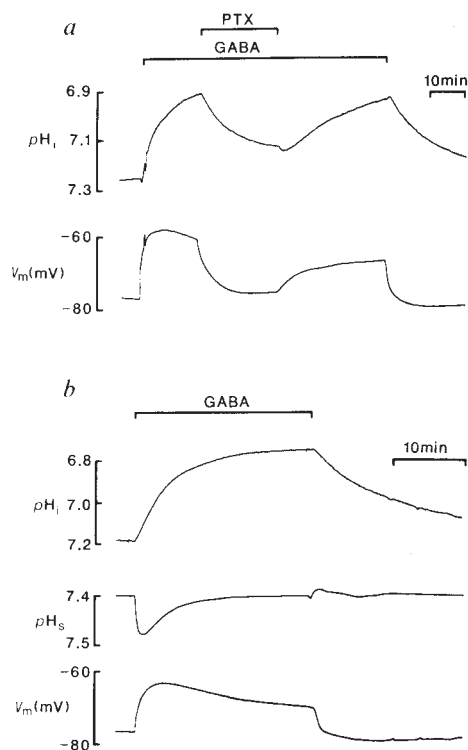


Fig. 4 Influence of GABA (10^{-4} M) on surface pH (pH_s) in pH_i and surface pH_s in the standard solution equilibrated with 5% CO_2 + 95% air. **a**, The GABA-induced intracellular acidosis and the accompanying depolarization are inhibited by PTX (10^{-4} M); and **b**, the decrease in pH_i is paralleled by an alkalosis at the surface of the muscle fibre. The slow decline of the GABA-induced depolarization after its peak value is attributable to a negative shift in $E_{HCO_3^-}$ caused by the simultaneous fall in pH_i . GABA had little effect on pH_i and pH_s in the absence of bicarbonate (not shown; compare Fig. 1). Surface pH was measured with a liquid-sensor pH -selective microelectrode with a shape similar to that of a patch pipette.

chloride-free solution equilibrated with 5% CO_2 and 95% air, and in a solution devoid of both chloride and bicarbonate, pH_i being measured simultaneously. In the standard CO_2 -free solution, GABA produced an increase in membrane conductance similar to that reported earlier³. After a 30-minute period of application of a solution devoid of both Cl^- and HCO_3^- , GABA (5×10^{-4} M) produced a very small increase in conductance that was only 1–3% of that seen under control conditions. In contrast to this, after a 30-minute application of the HCO_3^- -containing, Cl^- -free solution, GABA (5×10^{-4} M) induced a marked, PTX-sensitive increase in conductance, amounting to 13–17% of that seen under control conditions. Assuming constant-field conditions and that chloride, when present, is at equilibrium with resting V_m , the conductance data in Fig. 3 yield a mean estimate of the ratio $P_{Cl}:P_{HCO_3^-}$ of 1:0.43, when comparing results

obtained in the Cl^- -free, bicarbonate-containing solution with those made in the standard CO_2 -free solution. In two other experiments, the estimates for $P_{Cl}:P_{HCO_3^-}$ were 1:0.34 and 1:0.52. These values are in reasonable agreement with those obtained from measurements of the reversal potential.

A series of concentration-effect experiments showed that PTX had a more potent blocking action on the GABA-induced current carried by HCO_3^- than on the current mediated by Cl^- . A 50% block was induced by $5.8 \pm 0.3 \mu M$ PTX ($\pm s.e.m.$, $n = 4$) in the Cl^- -containing, HCO_3^- -free solution and by $1.3 \pm 0.1 \mu M$ PTX in the HCO_3^- -containing, Cl^- -free solution. The corresponding values for a 95% block were $138 \pm 11 \mu M$ and $53 \pm 4.3 \mu M$. This difference is not unexpected after previous observations¹⁵ and indicates a competitive interaction between PTX and permeant anions.

The above results indicate that GABA activates a conductance pathway for bicarbonate, which, considering the actions of PTX, is probably identical to that mediating Cl^- -currents. Further evidence in favour of this conclusion is that the fall in pH_i seen upon application of GABA in a HCO_3^- -containing solution was inhibited by PTX (Fig. 4a). Figure 4b shows that such a decrease in pH_i is linked to a surface alkalosis. This is consistent with an efflux of HCO_3^- and eliminates the possibility that application of GABA somehow leads to *de novo* generation of intracellular acid, which would be expected to produce a surface acidosis due to extrusion of acid.

Finally, we compared the effects of 10^{-4} M GABA on pH_i in solutions equilibrated with 1% or 5% CO_2 . In three experiments, the fall in steady-state pH_i ranged from 0.22–0.26 units at 1% CO_2 and from 0.41–0.44 units at 5% CO_2 . It should be noted that, in a living crayfish, the HCO_3^- -concentration both within excitable cells and in blood can be twice as high as the corresponding intra- and extracellular levels in the present experiments at 1% CO_2 (ref. 20).

We show here that GABA can induce a large fall in postsynaptic, intracellular pH and that the mechanism underlying this acidosis is a GABA-activated HCO_3^- -current. A GABA-induced effect on intracellular pH is of much interest for several reasons. Firstly, in excitable cells, changes in pH_i are known to influence the electrical activity of the cell membrane by affecting the properties of ion channels¹⁰. Secondly, in both nerve and muscle cells, a displacement of pH_i induces changes in the concentrations of other intracellular ions: for instance, by acting on transmembrane Cl^-/HCO_3^- -exchange, it could lead to a change in the intracellular Cl^- -concentration^{5–8} and, consequently, to a change in E_{GABA} . It is also notable that a GABA-gated HCO_3^- -conductance has a direct effect on E_{GABA} (see Fig. 2 and ref. 21). All these considerations suggest that a fall in postsynaptic, intracellular pH may be a physiologically important consequence of GABA-mediated inhibition. This hypothesis is strengthened from the observation⁴ that in mouse spinal neurons, the $P_{HCO_3^-}/P_{Cl}$ ratio of the GABA-gated channels is rather high, 0.18.

This work was supported by grants from the Academy of Finland. We thank R. D. Vaughan-Jones for comments on the manuscript.

Received 13 August; accepted 1 October 1987.

- Krnjević, K. *Physiol. Rev.* **54**, 418–540 (1974).
- Siggins, G. R. & Gruol, D. R. in *Handbook of Physiology, Section I: The Nervous System* Vol. IV, 1–114 (ed. Bloom, F. E.) (American Physiological Society, Bethesda, Maryland, 1986).
- Takeuchi, A. & Takeuchi, N. *J. Physiol., Lond.* **191**, 575–590 (1967).
- Bormann, J., Hamill, O. P. & Sakmann, B. *J. Physiol., Lond.* **385**, 243–286 (1987).
- Thomas, R. C. *J. Physiol., Lond.* **273**, 317–338 (1977).
- Moody, W. J. *J. Physiol., Lond.* **316**, 293–308 (1981).
- Moser, H. *J. Physiol., Lond.* **362**, 23–38 (1985).
- Galler, S. & Moser, H. *J. Physiol., Lond.* **374**, 137–151 (1986).
- Chesler, M. *J. Physiol., Lond.* **381**, 241–261 (1986).
- Moody, W. J. *A. Rev. Neurosci.* **7**, 257–278 (1984).
- Dudel, J., Finger, W. & Stettmeier, H. *Pflügers Arch.* **387**, 143–151 (1980).

- Zacher, J., Zacharová, D. & Hencsek, M. *Physiol. Bohemoslov.* **13**, 129–136 (1964).
- Dudel, J. & Rüdel, R. *Pflügers Arch.* **308**, 291–314 (1969).
- Dudel, J. *Pflügers Arch.* **371**, 167–174 (1977).
- Takeuchi, A. & Takeuchi, N. *J. Physiol., Lond.* **205**, 377–391 (1969).
- Adrian, R. H., Chandler, W. K. & Hodgkin, A. L. *J. Physiol., Lond.* **208**, 607–644 (1970).
- Constanti, A. & Smart, T. G. *Proc. R. Soc. B* **215**, 343–364 (1982).
- Goldman, D. E. *J. gen. Physiol.* **27**, 37–60 (1943).
- Hodgkin, A. L. & Katz, B. *J. Physiol., Lond.* **108**, 37–77 (1949).
- Gaillard, S. & Malan, A. *Molec. Physiol.* **4**, 231–243 (1983).
- Kelly, J. S., Krnjević, K., Morris, M. E. & Yim, G. K. W. *Expl Brain Res.* **7**, 11–31 (1969).
- Dudel, J. & Kuffler, S. W. *J. Physiol., Lond.* **155**, 514–529 (1961).
- Kaila, K. & Voipio, J. *J. Physiol., Lond.* **369**, 8P (1985).
- Kaila, K. & Vaughan-Jones, R. D. *J. Physiol., Lond.* **390**, 93–118 (1987).
- Vaughan-Jones, R. D. *J. Physiol., Lond.* **295**, 83–109 (1979).
- Vaughan-Jones, R. D. & Kaila, K. *Pflügers Arch.* **406**, 641–644 (1986).

Centrifugal fibres synapse on dopaminergic interplexiform cells in the teleost retina

Charles L. Zucker* & John E. Dowling

The Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138, USA

In teleost fish, centrifugal fibres originating in the olfactory bulb and containing FMRFamide-like and luteinizing hormone releasing hormone (LHRH)-like peptides project to the retina and terminate along the border of the inner nuclear and inner plexiform layers¹⁻³. Using a novel simultaneous two-colour immunolabelling technique, we have found that these centrifugal fibres are often closely apposed to the dopaminergic interplexiform cells. Contacts between centrifugal fibres and dopaminergic interplexiform cells were observed by electron microscopy to be conventional type synaptic junctions. Since the dopaminergic interplexiform cells make synapses on horizontal and bipolar cells, providing an intraretinal centrifugal pathway for information flow from the inner to the outer plexiform layers⁴⁻⁶, we conclude that every neuron in the teleost retina is potentially susceptible to central influences via these centrifugal fibres and dopaminergic interplexiform cells.

For light microscopy, centrifugal fibres in the retina of the white perch (*Roccus americanus*), and their cell bodies in the brain, were labelled with rabbit antisera directed to FMRFamide or [Trp⁷, Leu⁸]LHRH. Dopaminergic cells and their processes were labelled with a mouse monoclonal antibody directed to tyrosine hydroxylase (TOH; the rate limiting biosynthetic enzyme of dopamine). These antisera (FMRFamide at 1:1,000; TOH at 1:10,000) were applied simultaneously to 8- μ m-thick cryostat sections of 4% paraformaldehyde fixed tissue mounted on gelatin coated slides. The peptide antisera were visualized using a fluorescein isothiocyanate (FITC) labelled goat anti-rabbit IgG secondary antibody (Boehringer Mannheim) while the TOH was visualized with phycoerythrin-labelled goat anti-mouse IgG (Biomed). Both fluorochromes excite with blue light (450-490 nm) but have discrete emissions in the green and orange-red (long pass 520 nm). Using co-excitation of the two fluorochromes has allowed us to observe directly juxtapositions between centrifugal fibre varicosities with the cell bodies and processes of dopaminergic cells. Seeing both markers simultaneously greatly facilitates the screening of tissue for the relationships between the two.

For electron microscopy, a polyclonal rabbit anti-TOH antiserum (Eugene Tech, 1012) was used along with a peroxidase anti-peroxidase method that allows complete penetration of immunological reagents throughout a tissue block⁶. Following embedding in epon 812 and partial polymerization, 80- μ m-thick sections were taken and mounted on glass slides. Selected sections containing a dopaminergic cell body were remounted with a cyanoacrylate adhesive onto blank epon blocks and thin sectioned. The unique morphological appearance of the stained elements was utilized to distinguish processes of centrifugal fibres from those of the dopaminergic cells. The former contained large (~100 nm), highly immunoreactive, large dense-core vesicles as well as small electron-lucent vesicles, whereas the latter had no large dense-core vesicles.

Centrifugal fibres originating from the nucleus olfactoryretinalis run along the border of the inner plexiform and the inner nuclear layers. Fluorescence microscopy using our double labelling technique showed extensive clusters of FMRFamide-like

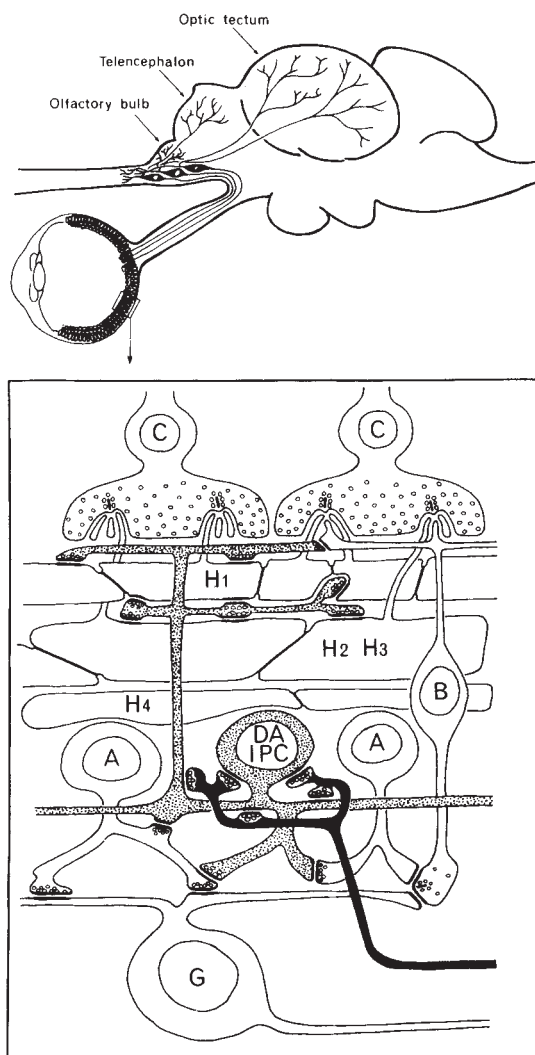


Fig. 1 A summary of the origins and synaptic connections of the olfactoryretinalis-dopaminergic interplexiform cell (DA-IPC) system in the white perch. The cell bodies of the olfactoryretinalis are located bilaterally near the midline at the base of the olfactory bulb where they send out processes toward the olfactory epithelium and bulb, as well as the telencephalon and optic tectum. Major axons decussate to the contralateral side where they enter the optic nerve and project out to the retina. In the retina, centrifugal fibres (solid black) ramify in the most distal inner plexiform layer and make extensive synaptic contacts onto the perikarya and proximal dendrites of dopaminergic interplexiform cells (stippled). Dopaminergic interplexiform cells, in turn, provide extensive synaptic output to cone horizontal cells (H1-H3) in the outer plexiform layer and some synaptic input to bipolar (B) cell dendrites in the outer plexiform layer and to amacrine (A) cells in the inner plexiform layer^{5,6}. The interplexiform cells themselves also receive input from amacrine cells, and some amacrine cells are probably contacted by centrifugal fibres¹². (C, cones; H4, rod horizontal cell; G, ganglion cell).

immunoreactive fibres and varicosities juxtaposed to dopaminergic cell bodies (Fig. 3a). Additionally, centrifugal fibres surround the proximal processes of dopaminergic cells for about 50-75 μ m from the cell body (Fig. 3b). In the goldfish retina, most if not all of the dopaminergic cells appear to be interplexiform cells⁵, and our observations indicate that this is also the case in the white perch. More than 700 dopaminergic cells (in both central and peripheral retinal regions) were observed adjacent to peptide-containing centrifugal fibres. We conclude, therefore, that the dopaminergic cells and processes contacted by the centrifugal fibres are principally, if not exclus-

* Present address: Neuroscience Unit, Eye Research Institute, 20 Staniford Street, Boston, Massachusetts 02114, USA.

Fig. 2 Electron micrographs of the white perch retina double labelled using a PAP technique for FMRFamide and TOH. Labelled centrifugal processes are easily distinguished from dopaminergic interplexiform cells since the former contain many highly immunoreactive large dense-cored vesicles. *a*, Cross-section through the centre of a dopaminergic interplexiform perikaryon (DA-IPC) with a centrifugal process (*) contacting it at its most proximal border. *b*, At higher magnification and through several serial sections, this same process is seen making a conventional synaptic contact onto the cell body. *c*, A proximal dendrite (note characteristic organelles) from the same dopaminergic interplexiform cell is shown receiving a synaptic contact from a centrifugal fibre. *d*, A centrifugal process is seen synapsing onto an interplexiform cell dendrite located about 25 μm from the same cell body. In all of the synapses shown, note the distal location of the immunoreactive large dense-cored vesicles.

Calibration bars: 5 μm (*a*); 0.5 μm (*b-d*).

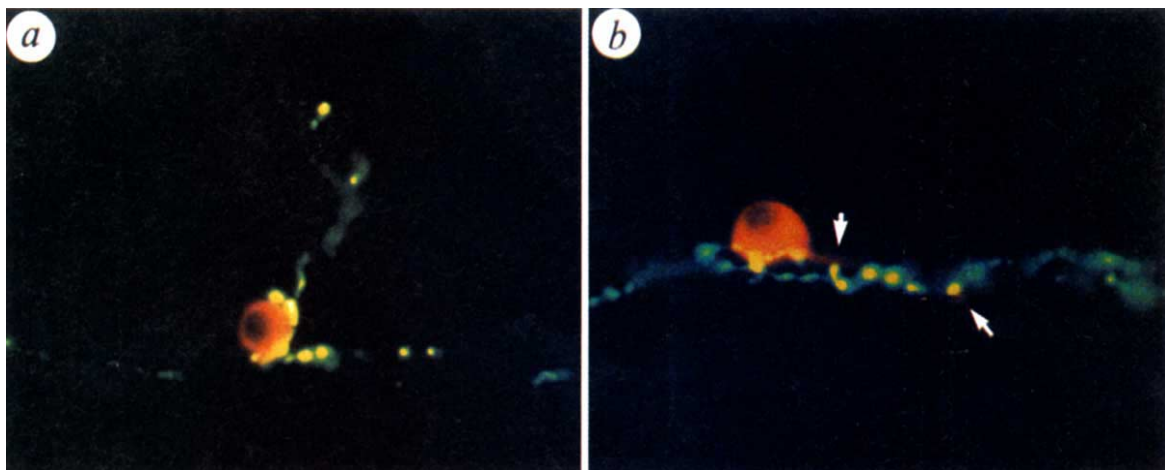
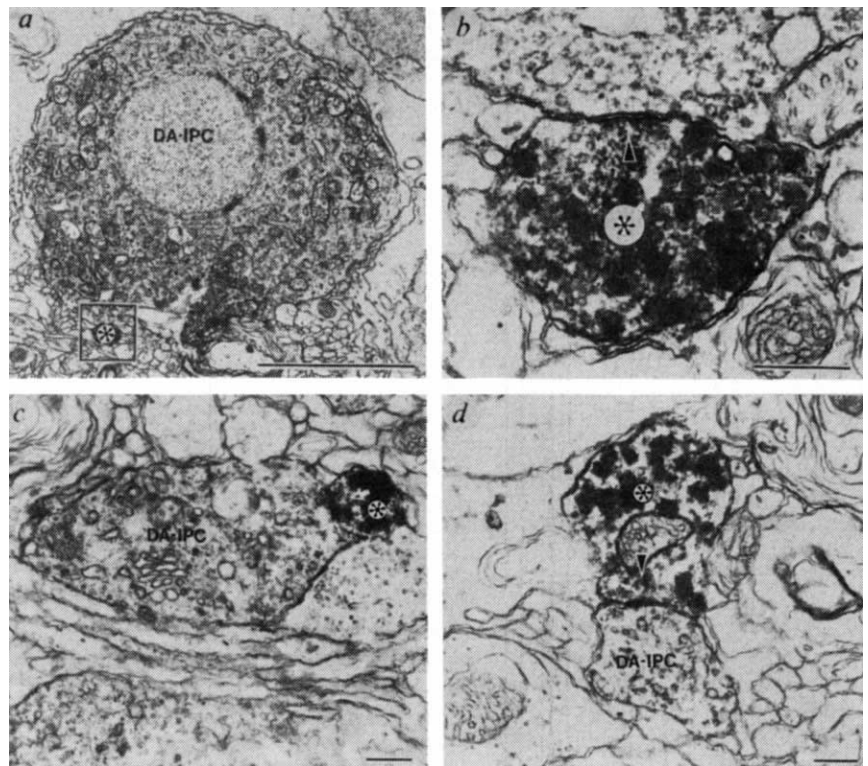


Fig. 3 Fluorescent micrographs of cryostat sections from white perch retina processed for simultaneous visualization of FMRFamide immunoreactive centrifugal fibres and TOH immunoreactive dopaminergic interplexiform cells. FITC- and phycoerythrin-labelled secondary antibodies were used; under blue excitation they resulted in apple-green centrifugal fibres and pumpkin-orange dopaminergic interplexiform cells and processes. Kodak Ektachrome P800/1600 was processed at push 1. *a*, Dopaminergic interplexiform cell perikaryon surrounded by a beaded string of centrifugal fibre varicosities. An extension of the centrifugal fibre is seen rising toward the outer plexiform layer. *b*, Dopaminergic interplexiform cell receiving extensive centrifugal contacts on its perikaryon and proximal dendrites. Both perikarya are about 12 μm in diameter.

ively, interplexiform cells. Centrifugal fibres occasionally ascend towards the outer plexiform layer. These ascending fibres invariably occur at the location of a dopaminergic interplexiform cell body, adjacent to where it extends a process towards the outer plexiform layer (Fig. 3*a*).

When visualized by electron microscopy, FMRF-amide-like immunoreactive processes showed both large dense-cored and small electron-lucent vesicles. These processes make conventional-type synapses, characterized by a cluster of small vesicles near the thickened presynaptic membrane (Fig. 2*b,c,d*). The large dense-cored vesicles tended to be distal to the presumed presynaptic site. When stained with anti-FMRFamide or with anti-[Trp⁷,Leu⁸]LHRH all large dense-cored vesicles were immunoreactive, whereas the small electron-lucent vesicles were

always unstained. This indicates that both peptides co-localize within the large vesicles leaving a probable third neuroactive substance unidentified in the small vesicles.

Junctions between the centrifugal fibres and the cell perikarya and proximal processes of dopaminergic cells were frequently observed in double-labelled tissue, confirming the fluorescence microscopy observations. Figure 2*a* shows a section through a TOH-labelled dopaminergic cell perikaryon. A labelled centrifugal process is seen along its proximal border. Serial sections (Fig. 2*b*) revealed several synaptic junctions on the perikaryon. Additional centrifugal input was observed on a process from this same dopaminergic cell that was followed by serial sections (Fig. 2*c,d*). At least 13 conventional input synapses from centrifugal fibre terminals were observed in semi-serial sections

through half of this cell body and its proximal dendritic tree. Centrifugal fibres did not appear to interact with dopaminergic interplexiform cells further than 75 μm from the cell body or into the proximal inner plexiform layer.

We also studied the distribution of FMRFamide and LHRH elsewhere in the white perch brain. Using either anti-FMRFamide or anti-[Trp⁷,Leu⁸]LHRH, bilateral clusters of large (25–35 μm diameter) cell bodies (35–50 per side) were visualized centred about the midline at the base of the olfactory bulb near the ventral surface. These clusters showed immunoreactivity for both peptides, but they were the only cells stained with the FMRFamide antibody in the white perch brain, indicating that they are likely to be the cells of origin of the centrifugal fibres. These cells send a profusion of processes throughout the olfactory bulb, telencephalon and optic tectum, and are themselves surrounded by varicosities, which are immunoreactive to FMRFamide. It appeared as if a single main axonal fibre projects from each cell body to enter the optic nerve.

We conclude that centrifugal fibres, originating from the olfactory system, interact with the retinal dopaminergic interplexiform cell system. This circuitry, summarized in Fig. 1, indicates that interplexiform cells represent the most distal reaches of the centrifugal fibre system in the fish retina. Thus, centrifugal influences can reach bipolar and horizontal cells via dopaminergic interplexiform cells and even the photoreceptors via feedback from cone horizontal cells⁷. It follows that every retinal neuron is potentially susceptible to central influences in the teleost. The physiological significance of this central input to the dopaminergic interplexiform cells is not clear. Present evidence suggests that the interplexiform pathway modulates

light responsiveness and receptive field size of horizontal cells in the teleost outer plexiform layer^{8–11}.

We also show that centrifugal terminals within the retina probably contain two peptides that co-localize in the same large dense-cored vesicles. In these terminals, small electron-lucent vesicles are also observed that do not appear to be immunoreactive. Therefore, these centrifugal fibres may contain an as yet unidentified third neuroactive substance. The possibility of a third neuroactive substance being directly released at synapses onto dopaminergic interplexiform cells, in addition to the two peptides which have already been identified, suggests that the interaction between the centrifugal fibres and dopaminergic interplexiform cells is a complex one.

We thank T. O'Donohue, NIH, for providing FMRFamide, N. Sherwood, University of Victoria, for [Trp⁷, Leu⁸]LHRH and R. Vulliet for TOH. This work was partially supported by grants from the NIH.

Received 4 August; accepted 6 October 1987.

1. Springer, A. D. *J. comp. Neurol.* **214**, 404–415 (1983).
2. Munz, H., Claas, B., Stumpf, W. E. & Jennes, L. *Cell Tissue Res.* **222**, 313–323 (1982).
3. Stell, W. K., Walker, S. E., Chohan, K. S. & Ball, A. K. *Proc. natn. Acad. Sci. U.S.A.* **81**, 940–944 (1984).
4. Dowling, J. E. & Ehinger, B. *Science* **188**, 270–273 (1975).
5. Dowling, J. E. & Ehinger, B. *Proc. R. Soc. B201*, 7–26 (1978).
6. Eldred, W. D., Zucker, C. L., Karten, H. J. & Yazulla, S. J. *Histochem. Cytochem.* **31**, 285–292 (1982).
7. Baylor, D. A., Fuortes, M. G. F. & O'Bryan, P. M. *J. Physiol., Lond.* **214**, 256–294 (1971).
8. Hedden, W. L. & Dowling, J. E. *Proc. R. Soc. B* **201**, 27–55 (1978).
9. Teranishi, T., Negishi, K. & Kato, S. *J. Neurosci.* **4**, 1271–1280 (1984).
10. Piccolino, M., Neyton, J. & Gerschenfeld, H. M. *J. Neurosci.* **4**, 2477–2488 (1984).
11. Mangel, S. C. & Dowling, J. E. *Proc. R. Soc. B231*, 91–121 (1987).
12. Dowling, J. E. & Cowan, W. M. Z. *Zellforsch. mikrosk. Anat.* **71**, 14–28 (1966).

Non-responsiveness to a foot-and-mouth disease virus peptide overcome by addition of foreign helper T-cell determinants

Michael J. Francis*, Gillian Z. Hastings*,
Andrew D. Syred*, Brian McGinn†,
Fred Brown* & David J. Rowlands*

* Wellcome Biotechnology, and † Wellcome Foundation,
Beckenham, Kent BR3 3BS, UK

Study of the immune response to synthetic antigens has shown that uncoupled peptides can realize their potential as vaccines only if they contain domains that react with helper T-cell receptors and Ia antigens^{1,2} in addition to antibody binding sites. Here we consider whether genetically restricted non-responsiveness to an uncoupled peptide could be overcome by synthesizing a peptide with an additional helper T-cell epitope from a different protein. We demonstrate that H-2^d mice, which are non-responders to the 141–160 VP1 peptide of foot-and-mouth disease virus (FMDV), can be converted into responders by immunization with peptides containing the FMDV sequence with defined 'foreign' helper T-cell determinants from ovalbumin or sperm whale myoglobin. Furthermore, the virus-neutralizing activity of the antibody raised against peptide was dependent on the determinant used. Thus, FMDV peptides with the added sequences 323–339 from ovalbumin and 132–148 from sperm-whale myoglobin elicited a high degree of neutralizing activity in B10.D2 mice. The sera from mice which received the peptide with the added sequence 105–121 from sperm whale myoglobin did not neutralize the virus, although they had high levels of anti-141–160 FMDV peptide activity. Our data indicate that the T-cell help given by the 'foreign' epitopes is B-cell clone specific. These results are likely to have important implications for the design of peptide vaccines.

A synthetic peptide comprising the 141–160 region of VP1 from FMDV elicits a neutralizing antibody response that protects guinea pigs against experimental infection^{3,4}. The response to this peptide is not carrier-restricted^{5,6}, and the uncoupled peptide is immunogenic when presented in liposomes^{6,7} or in incomplete Freund's adjuvant^{7,8}, provided that a non-natural cysteine is added to the carboxy terminus. These results led us to conclude⁶ that this peptide not only contained an important site(s) that could elicit neutralizing antibodies against the virus (B-cell epitope), but also a site(s) capable of inducing helper T-cell activity⁹ for this response (Th-cell epitope). This conclusion is supported by the stimulation of T-cell proliferation by the amino-acid sequence 150–160 of the peptide *in vitro*^{7,8}. However, as this peptide can only contain a few Th-cell epitopes problems could arise in obtaining good immunity within an outbred population or in other species⁸.

Indeed, the response of potential vaccine target species (cattle and pigs) to this free peptide is not as good as in guinea pigs (M.J.F., unpublished data), although the antibodies elicited by the peptide are qualitatively similar in all three species (E. J. Ouldrige, N. R. Parry & M.J.F., unpublished data). To overcome this genetic restriction of the immune response to the free peptide, we have investigated whether such non-responsiveness in a laboratory model could be modified by using a peptide with additional Th-cell epitopes. Control of FMD requires regular vaccination at intervals of 4–6 months to maintain large amounts of circulating antibody in the susceptible population, since natural B-cell or T-cell memory would be unable to provide protection in the event of severe virus challenge. Therefore, we considered the possibility of using 'foreign' Th-cell epitopes from non-FMDV proteins to modify the response to the FMDV peptide. Recently it has been shown that non-immunogenic B cell epitopes can be made immunogenic by conjugation with a 'natural'¹⁰ or by co-polymerization with a 'foreign'¹¹ T cell epitope.

It was first necessary to determine whether the response to the 141–160 peptide was restricted in a genetically defined population of inbred mice. A range of congenic B10 mice,

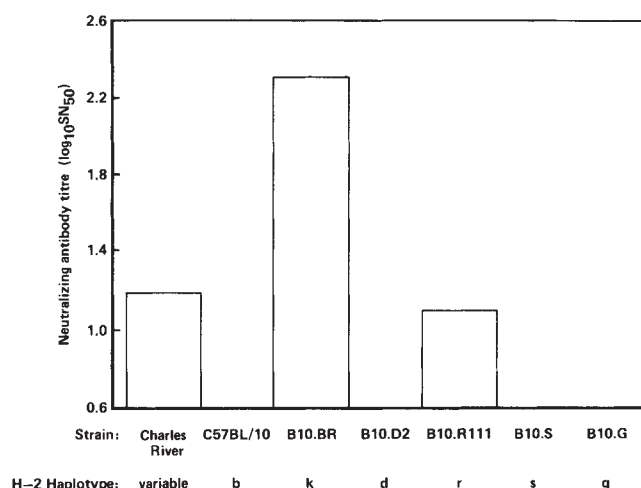


Fig. 1 Mean neutralizing antibody response of outbred and congenic mice to uncoupled 141-160 peptide. Groups of 10 outbred Charles River and congenic B10 mice were inoculated intramuscularly with 100 µg (equivalent to 43 nM) of uncoupled 141-160 Cys Gly peptide emulsified in incomplete Freund's adjuvant. Serum samples were collected at 56 d and tested for the presence of neutralizing antibody against 100TCID₅₀ of virus in a micro-neutralization test using IBRS2 cells²².

genetically identical at all loci except those of the *H-2* complex, were therefore immunized with 141-160-Cys-Gly peptide and their neutralizing antibody responses compared with those of an outbred population (Fig. 1). Only two of the congenic strains studied, B10.BR and B10.R111 (71NS), responded to the free peptide: The B10.BR strain, which is *H-2^k* haplotype, gave a particularly high response in both its neutralizing and anti-peptide antibody which was significantly higher than that of outbred Charles River mice. None of the other congenic strains studied produced a detectable neutralizing antibody response to the free 141-160 peptide. Similar MHC-restriction of the immune response to the FMDV 141-160 sequence has also been observed in congenic mice with a BALB background.

Among the non-responding congenic mice was the B10.D2 strain, which is *H-2^d* haplotype. As several *H-2^d*-restricted Th-cell epitopes have been characterized^{12,13}, this strain was chosen for our study. From the published epitopes we selected three, one from ovalbumin contained within the sequence 323-339 (OVA) (ref. 14), and two from sperm-whale myoglobin contained within the sequences 132-148 (SWMI) (refs 15, 16) and 105-121 (SWMII) (refs 17, 18). Peptides were synthesized¹⁹ with the FMDV VP1 141-160 sequence, followed at the carboxy terminus by 17 additional amino-acid residues containing one of the foreign T-cell sites. As a control, a fourth peptide of similar length was synthesized to cover amino acids 141-160 and 161-177 of FMDV VP1. Although this peptide provided the best negative control for the non-responding strains, by adding an extra 17 natural residues there was a risk of adding further natural T-cell epitopes, and thus overcoming the non-responsiveness to the sequence 141-160 alone. Each peptide also had a non-natural cysteine residue added to the carboxy terminus for increased immunogenicity^{7,8}. In addition to the B10.BR and B10.D2 congenic strains, we examined the response of another inbred strain of mice (BALB/c), also *H-2^d* haplotype and a poor responder to the 141-160 peptide, to determine whether the Th-cell mediated non-responsiveness was *H-2* restricted and whether this could be overcome in genetically distinct populations of a similar *H-2* haplotype.

As expected, the B10 BR (*H-2^k*) mice produced an anti-peptide antibody response to the 141-160 peptide after inoculation with 141-177 Cys FMDV peptide or any of the three peptides containing the 141-160 sequence from FMDV and an additional foreign Th-cell epitope (Fig. 2a). This was presum-

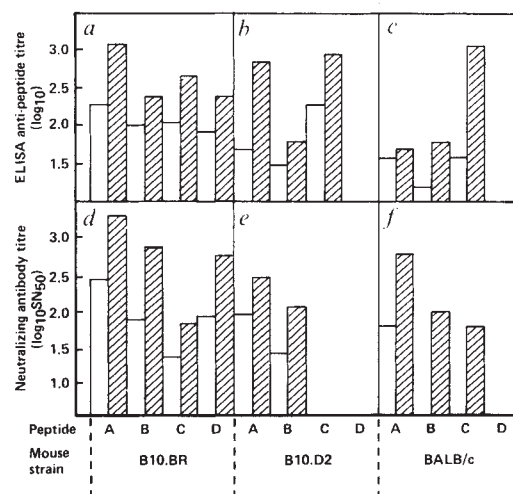


Fig. 2 Anti-peptide and neutralizing antibody response of mice to FMDV peptides with foreign T-cell epitopes. Groups of 8 inbred B10.BR, B10.D2 and BALB/c mice were inoculated intramuscularly with 25 nM (equivalent to approximately 100 µg) of FMDV VP1 141-160 peptide plus foreign T-cell epitopes, referred to as OVA (A), SWMI (B) and SWMII (C) or FMDV 161-177 sequence (D), in incomplete Freund's adjuvant. Each mouse was bled before inoculation and at 28 d post-inoculation. All mice were reinoculated as above at 63 d, and further serum samples were collected 28 d later. The 28 d post-primary inoculation (open columns) and 28 d post-secondary inoculation (hatched columns) serum samples were subsequently analysed for anti-peptide activity (a-c) in an indirect ELISA (ref. 8) and virus-neutralizing activity (d-f) against a 100TCID₅₀ dose of virus in a micro-neutralization assay²².

ably caused by helper T-cell activity provided by a Th-cell epitope in the 141-160 part of the sequence. In contrast, the B10.D2 and BALB-c (*H-2^d*) mice both failed to produce an antibody response to the 141-160 peptide after immunization with 141-177 Cys (Fig. 2b and c). Therefore, the addition of 17 natural amino-acid residues to the carboxy-terminus did not overcome the non-responsiveness of these two strains of mice to uncoupled FMDV peptide, but each of the peptides with a foreign Th-cell epitope elicited an antibody response in these mice to the 141-160 peptide (Fig. 2b and c). Thus we have shown that non-responsiveness of *H-2^d* haplotype mice to the 141-160 peptide can be overcome by adding a peptide sequence with a foreign Th-cell epitope at its carboxy terminus.

We next determined whether these anti-peptide antibodies were also functional antibodies with virus-neutralizing activity. The analysis (Fig. 2d-f) reveals some interesting but unexpected information on the Th-cell control of B-cell antibody production. Neutralizing antibody activity was detected in the B10.BR strain of mice regardless of the peptide received and no neutralizing antibody was detected in the B10.D2 or BALB/c mice which received the FMDV sequence peptide (141-177). Neutralizing antibodies were evoked in these *H-2^d* mouse strains by the peptides containing OVA or SWMI Th-cell epitopes, but the 141-160 peptide with the SWMII Th-cell epitope only produced a neutralizing antibody response after reinoculation in the BALB/c strain. Therefore, although the foreign Th-cell epitope SWMII induced an anti-141-160 peptide response in B10.D2 mice, these antibodies did not have virus-neutralizing activity, despite the production of neutralizing antibodies by this peptide in B10.BR mice, indicating that the conformation of the 141-160 sequence is not significantly affected by the addition of the SWMII epitope. Furthermore, the antisera with anti-peptide and neutralizing activity also reacted with virus in both indirect and sandwich enzyme-linked immunosorbent assays (ELISA) and the antisera with anti-peptide activity alone did not (Table 1). This result indicates that Th-cell epitopes could control antibody production of specific B-cell clones, which might be

Table 1 Anti-virus activity measured by enzyme-linked immunosorbent assay

Mouse strain	Peptide	Anti-virus ELISA (log ₁₀)	
		Indirect	Sandwich
B10.BR	A	2.3	1.8
	B	2.2	2.0
	C	1.9	1.9
	D	2.0	2.0
B10.D2	A	1.8	1.7
	B	1.3	1.4
	C	<1.0	<1.0
	D	<1.0	<1.0
BALB/c	A	1.5	1.3
	B	1.5	1.4
	C	1.9	1.3
	D	<1.0	<1.0

Sera were collected from B10.BR, B10.D2 and BALB/c mice 28 days after a reinoculation with peptides A (141–160 OVA), B (141–160 SWMI), C (141–160 SWMII) and D (141–177). Anti-virus activity was assessed in an indirect ELISA, in which purified virus is bound directly to a solid phase, and a sandwich ELISA, in which purified virus is bound to a solid phase by a capture antibody²¹. Titres were calculated by reference to a negative standard (a 1:10 dilution of pre-inoculation serum from the same strain of mouse).

a feature of the T-cell epitope itself or of its position in relation to the B-cell site it is regulating. Such specificity of control has been described for helper T-cell determinants from hepatitis B virus²⁰. This result also suggests that co-polymerization of T- and B-cell epitopes with glutaraldehyde¹¹ is unlikely to produce an optimal construction and that a simple anti-peptide test alone can provide misleading data.

To characterize the nature of the virus-neutralizing and non-neutralizing antibody populations, both of which recognize the 141–160 peptide in an indirect ELISA (Fig. 2b and e), we tested the 28 days post-reinoculation sera from the B10.D2 mice that received the 141–160 peptide with the OVA or SWMII T cell sites against a range of peptides located within the 141–160 region. The results of this analysis (Table 2) show that the peptide with the SWMII site produces antibodies that only recognize peptides from the amino terminus. They do not react with any peptides from the 150–160 region. In contrast, the antibodies to the peptide with the OVA T-cell site, which neutralize the virus, recognize residues from the 147–156 region. This portion of the peptide has been shown to be critical for the induction of virus-neutralizing antibodies⁸.

In conclusion, we show here that genetically controlled non-responsiveness to FMDV 141–160 peptide in mice can be overcome by adding foreign Th-cell epitopes, and that this helper

activity is B-cell clone specific. In addition, these experiments demonstrate that a carrier protein will be sufficient if it provides an appropriate helper T-cell epitope and that size *per se* is not a factor. The significance of this result for FMD vaccination is that the poor response to uncoupled FMDV peptide in cattle may be overcome by adding further cattle Th-cell epitopes to the sequence. Furthermore, this FMDV mouse model is a powerful method for elucidation of the interaction between the T- and B-cell epitopes on single peptide molecules and whether the addition of further T-cell epitopes, or the manipulation of existing epitopes, can modify MHC restriction of the immune response. Until such studies are completed we cannot predict whether single or mixed peptide preparations will be as effective as native antigens in outbred populations, but a fuller understanding of the immune response will enable synthetic immunogens to be designed with known B-cell and T-cell epitopes that could overcome much of the variation within and between species encountered with current vaccine preparations.

Received 13 July; accepted 23 September 1987.

- Rosenthal, A. S. *Immun. Rev.* **40**, 136–152 (1978).
- Babbitt, B. *et al. Nature* **317**, 359–361 (1985).
- Bittle, J. L. *et al. Nature* **298**, 30–37 (1982).
- Pfaff, E. *et al. EMBO J.* **1**, 869–874 (1982).
- Francis, M. J. *et al. in Vaccines 85* (eds Lerner, R. A., Chanock, R. M. & Brown, F.) 204–210 (Cold Spring Harbor, New York, 1985).
- Francis, M. J. *et al. J. gen. Virol.* **66**, 2347–2354 (1985).
- Francis, M. J. *et al. in Vaccines 87* (eds Chanock, R. M., Ginsberg, H., Lerner, R. A. & Brown, F.) 60–67 (Cold Spring Harbor, New York, 1987).
- Francis, M. J. *et al. Immunology* **60**, 1–6 (1987).
- Mitchison, N. A. *Eur. J. Immun.* **1**, 18–27 (1971).
- Good, M. F. *et al. Science* **235**, 1059–1062 (1987).
- Leclerc, C. *et al. Eur. J. Immun.* **17**, 269–273 (1987).
- DeList, C. & Berzofsky, J. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7048–7052 (1985).
- Rothbard, J. B. *Ann. Inst. Pasteur* **137E**, 518–528 (1986).
- Shimonkevitz, R. *et al. J. Immun.* **133**, 2067–2074 (1984).
- Berkower, I. *et al. J. Immun.* **132**, 1370–1378 (1984).
- Berkower, I. *et al. J. Immun.* **135**, 2628–2634 (1985).
- Streicher, H. Z. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 6831–6835 (1984).
- Berkower, I., Buckenmeyer, G. K. & Berzofsky, J. A. *J. Immun.* **136**, 2498–2503 (1986).
- Houghten, R. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5131–5135 (1985).
- Milich, D. R., McLachlan, A. & Thornton, G. B. *in Vaccines 87* (eds Chanock, R. M., Ginsberg, H., Lerner, R. A. & Brown, F.) 50–55 (Cold Spring Harbor, New York, 1987).
- Ouldridge, E. J., Barnett, P. V. & Rweyemamu, M. M. *in Curr. Top. Vet. Med. Anim. Sci.* (eds Wardley, R. C. & Crowther, J. R.) Vol. 22, 142–146 (Nijhoff, The Hague, 1984).
- Francis, M. J. & Black, L. *J. Hyg. Camb.* **91**, 329–334 (1983).

Expression and function of the CD3-antigen receptor on murine CD4⁺8⁺ thymocytes

Wendy L. Havran*, Martin Poenie*, John Kimura*, Roger Tsien*, Arthur Weiss† & James P. Allison*

* Cancer Research Laboratory, University of California, Berkeley, California 94720, USA

† Department of Medicine, Howard Hughes Medical Institute, University of California, San Francisco, California 94143, USA

Table 2 Nature of anti-peptide 141–160 response in B10.D2 mice

Peptide used in ELISA	Peptide used for inoculation	
	141–160 OVA	141–160 SWMII
141–160Cys	2.9*	3.0
141–153Cys	2.4	2.5
141–151Cys	1.4	2.5
141–149Cys	1.2	2.4
141–147Cys	<1.0	2.1
150–160Cys	1.6	<1.0
152–160Cys	1.6	<1.0
154–160Cys	1.1	<1.0
156–160Cys	<1.0	<1.0

Sera collected from B10.D2 mice 28 d after reinoculation with a 141–160 OVA or 141–160 SWMII peptides were examined in an indirect ELISA (ref. 8) against a range of peptides within the 141–160 region. Titres were calculated by reference to a negative standard (a 1:10 dilution of pre-inoculation serum from the same strain of mouse).

Mature T cells arise from progenitor cells by a complex and poorly understood process of differentiation in the thymus. Thymocytes can be divided into four major compartments on the basis of surface expression of the murine equivalents of CD8 (Lyt-2) and CD4 (L3T4) (refs 1, 2). Functionally mature thymocytes express only CD4 or CD8. The CD4⁺8[−] subset contains progenitor cells capable of giving rise to all the phenotypic and functional classes of T cells on adoptive transfer³. The function of the major population, the CD4⁺8⁺ cells, which carry both the CD4 and CD8 antigens, in thymic differentiation is controversial. It has been variously proposed that they represent terminally differentiated cells which die in the thymus¹ or that they represent an intermediate stage and can differentiate into functionally and phenotypically mature single positive T cells⁴. The CD3-antigen receptor com-

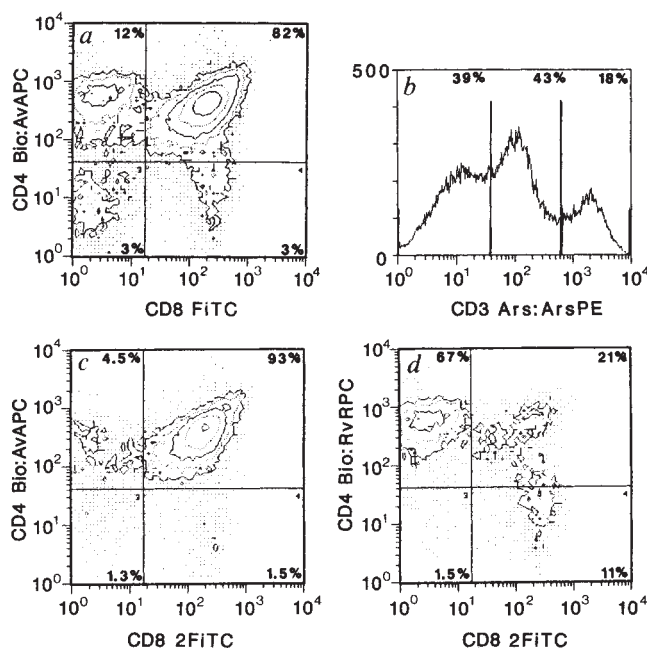


Fig. 1 Expression of CD3 by murine thymocytes. Thymocytes from a 12-week-old BALB/c female were stained with fluorescein isothiocyanate (FITC)-conjugated anti-CD8 (Lyt-2, 53-6.71), biotin-conjugated anti-CD4 (L3T4, GK1.5) and arsenyl (ARS)-derivatized anti-CD3 (hamster anti-murine CD3 λ , 500A2). Allophycocyanin (APC)-conjugated streptavidin and affinity-purified, phycoerythrin (PE)-conjugated rabbit anti-ARS were used as second reagents. Fluorochrome-conjugated monoclonal antibodies unreactive with murine thymocytes were used to control for nonspecific fluorescence. Samples were analysed by multiparameter flow cytometry using a FACS 440 system with an argon ion laser (488 nm) and a rhodamine dye laser (620 nm) (Beckton Dickinson). Data analysis was performed using the Consort 30 system (Beckton Dickinson). Panel *a*, Two-colour analysis of ungated thymocytes: vertical axis, CD4; horizontal axis, CD8. Panel *b*, histogram of anti-CD3 staining. Panel *c*, two-colour analysis of thymocytes gated for intermediate staining (38–650 units) with anti-CD3. Panel *d*, two-colour analysis of thymocytes gated for bright (651–9000 units) staining with anti-CD3. In all cases, the % total cells staining in each quadrant (*a*, *c*, *d*) or between markers (*b*) is indicated.

plex^{5–11} is probably important in thymic differentiation. The receptor has two functions: recognition and transmembrane signalling¹². To help clarify the function of CD4⁺CD8⁺ thymocytes in thymic differentiation, the expression and function of the antigen receptor complex on these cells should be determined. We show here that most CD4⁺CD8⁺ thymocytes express CD3 which functions in transmembrane signalling. The consequences of this signalling differ from those in mature T cells, however, in that the CD4⁺CD8⁺ cells do not secrete IL-2, express IL-2 receptor, or proliferate.

Three-colour immunofluorescence analysis was employed to simultaneously analyse the expression of CD3, CD4 and CD8 on murine thymocytes. Figure 1*a* represents the typical distribution of CD8 and CD4 on thymocytes. 80% of the cells expressed both CD8 and CD4, 3% did not express either, 12% were CD4 single positive, and 3% of the cells expressed CD8 alone. To assess CD3 expression we used 500A2, a hamster monoclonal antibody directed against the ϵ chain of murine CD3 (ref. 13). The histogram in Fig. 1*b* demonstrates binding of the anti-murine CD3 monoclonal antibody 500A2 to thymocytes. Thymocytes can be divided into CD3 negative, CD3 low-, and CD3 high-expressing cells. The majority of the cells showing high amounts of CD3 expressed only CD4 or CD8 (Fig. 1*d*). Of the cells with low expression of CD3, 93% were CD8⁺, CD4⁺ (Fig. 1*c*). In total 54% of CD4⁺CD8⁺ thymocytes in this experiment expressed significant amounts of CD3.

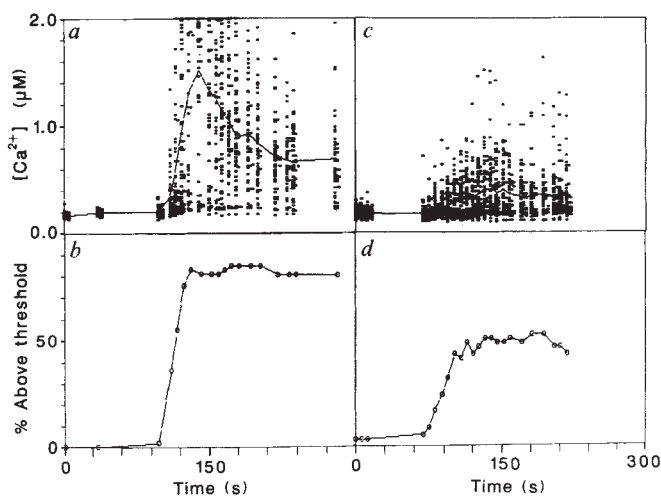


Fig. 2 $[Ca^{2+}]_i$ in individual CD4⁺CD8⁺ thymocytes. 10^6 cells ml⁻¹ were loaded with 1 μ M fura-2 acetoxyethyl ester in medium at 37°C for 30 min. Cells were then washed by brief centrifugation and resuspended in Hank's balanced salt solution supplemented with 5% fetal calf serum. Washed cells were settled onto coverslips coated with a 10-fold dilution of Cell-Tak. The resulting monolayer of cells was then used for the fluorescence ratio imaging measurements of $[Ca^{2+}]_i$ as described^{25,26}. The average $[Ca^{2+}]_i$ for individual cells in the field was determined by manually delimiting the cells on screen with a circle and using the computer to average the $[Ca^{2+}]_i$ for all the pixels within the circle. For graphs *a* and *c*, each cell is represented by a block which is positioned on the y-axis according to $[Ca^{2+}]_i$ and on the x-axis according to time. The mean $[Ca^{2+}]_i$ for all the cells in the field is also displayed as a solid line overlaid on the graph. The measurements begin with the unstimulated cells. At 90 s (panels *a*, *b*) or 60 s (panels *c*, *d*) anti-CD3 (500A2) was added. The subsequent rise in $[Ca^{2+}]_i$ for splenic T cells (panel *a*) or sorted CD4⁺CD8⁺ thymocytes (panel *c*) follows the addition of the antibody. The response was not increased by the addition of cross-linking antibody. An arbitrary threshold of 0.3 μ M was used to divide responsive from non-responsive cells, as effectively all of the cells in the unstimulated populations were below this level. Panels *b* (splenic T cells) and *d* (CD4⁺CD8⁺ thymocytes) show the % cells above this 0.3 μ M threshold during the course of the measurement. FACS analysis of the CD4⁺CD8⁺ cells used in the experiment presented in panels *c* and *d* indicated that 54% expressed CD3.

Studies using antibodies to the antigen-receptor complex as agonists have established that the antigen receptor mediates a transmembrane signal, which leads to hydrolysis of phosphatidylinositol 4,5-bisphosphate and the generation of intracellular second messengers, resulting in an increase of cytoplasmic free calcium ($[Ca^{2+}]_i$) and activation of protein kinase C (ref. 12). These events are causally linked to T cell activation. Previous studies have shown that CD4⁺CD8⁺ thymocytes are not functionally mature². It was therefore of interest to determine whether the CD3-antigen receptor complex expressed by CD4⁺, CD8⁺ thymocytes could mediate transmembrane signalling. We isolated CD4⁺CD8⁺ thymocytes by fluorescence-activated cell sorting (FACS) and analysed them for the early events of signal transduction. FACS analysis was performed simultaneously on cells from each thymus used in the calcium studies.

As the multicolour flow cytometric analysis indicated that only 50–60% of the CD4⁺CD8⁺ cells expressed detectable amounts of CD3, we used image analysis to assess cytoplasmic calcium concentration in individual cells. Figure 2*a* illustrates the response of individual panned splenic T cells to antibody against CD3. Most cells exhibited a very rapid response to the antibody, reaching a mean $[Ca^{2+}]_i$ of 1.5 μ M within 50 seconds of antibody addition. As shown in Fig. 2*b*, 88% of splenic T cells responded by increasing $[Ca^{2+}]_i$ above a threshold of 0.3 μ M. The same results were obtained if the T cells were incubated with the

Table 1 Induction of IL-2 secretion, IL-2 receptor expression and cell proliferation by antibody to CD3

Cells	Treatment*	Adherent cells	IL-2‡ (U ml ⁻¹)	IL-2R§ (MFI)	Proliferation (c.p.m.)
Splenic T ⁶	media	—	0.0	35	1,452
	anti-CD3	—	2.3	139	2,040
	anti-CD3	+	15.6	NT	63,460
	PMA + ionomycin	—	13.3	121	66,206
	anti-CD3 + 1U/ml IL-2	+	NT	NT	44,039
	anti-CD3 + 10U/ml IL-2	+	NT	NT	74,942
Splenic T + anti-CD4, 8*	media	—	0.0	NT	1,129
	anti-CD3	—	1.9	NT	1,161
	anti-CD3	+	14.9	NT	60,188
	PMA + ionomycin	—	14.1	NT	59,465
	anti-CD3 + 1U/ml IL-2	+	NT	NT	61,143
	anti-CD3 + 10U/ml IL-2	+	NT	NT	75,857
CD4 ⁺ 8 ⁺ thymus*	media	—	0.0	31	1,154
	anti-CD3	—	0.0	33	1,272
	anti-CD3	+	0.0	NT	2,301
	PMA + ionomycin	—	0.0	29**	4,046
	anti-CD3 + 1U/ml IL-2	+	NT	NT	2,057
	anti-CD3 + 10U/ml IL-2	+	NT	NT	2,851

The data presented here were obtained in a single representative experiment. NT, not tested.

* 5×10^4 T cells were cultured in media alone, a 1:4 dilution of 500 A2 anti-CD3 culture supernatant alone or with 1 or 10 units ml⁻¹ of rIL-2 (Cetus), or 10 ng ml⁻¹ PMA and 1 μ M ionomycin.

† Peritoneal exudate macrophages were obtained from a C57BL/6 mouse and X-irradiated for 20 minutes. 1×10^4 cells were added to cultures as indicated.

‡ Supernatants were collected after 24 h and IL-2 activity measured as relative uptake of 3-(4,5- α , methylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (Sigma) by the IL-2 dependent cell line CTLL-2 in a 24 h assay²⁴. Units of IL-2 were calculated by comparison to a standard (lot ISDP-841) provided by the NCI Biological Response Modifiers program. Units represented the mean of triplicate cultures. Standard errors were less than 10% of the mean. Similar results were obtained in 4 separate experiments.

§ IL-2 receptor expression was measured on cells cultured for 24 h with the indicated stimulus and then stained with biotin-conjugated anti-IL-2 receptor (PC61.5) followed by allophycocyanin-conjugated streptavidin. Samples were analysed as described in Fig. 1. Results are expressed as mean fluorescence intensity (MFI).

|| Thymidine incorporation was measured by culturing 5×10^4 T cells with the indicated treatment for 72 h in 96 well microtitre wells. 12 h before collection, 0.5 μ Ci ³H-thymidine was added to each well. Thymidine incorporation is presented as the mean c.p.m. of triplicate cultures. Standard errors were less than 10% of the mean. Results obtained in this experiment are representative of 6 separate experiments.

* T spleen cells were isolated after incubating C57BL/6 spleen cells on a 25 cm² tissue culture flask coated with affinity purified rabbit anti-mouse immunoglobulin for 1 h. Non-adherent cells were recovered and determined to be >95% Thy-1⁺ by FACS analysis.

* Isolated T spleen cells and thymocytes from a 4–6 week C57BL/6 mouse were incubated with FITC-anti-Lyt-2 (CD8) and biotin-anti-L3T4 (CD4) for 1 h at 4°C. Cells were washed and incubated an additional hour with PE-avidin. Lyt-2⁺L3T4⁺ (CD4⁺8⁺) thymocytes were collected after FACS sorting. Reanalysis of sorted cells demonstrated >95% purity. Staining of control cells and thymocytes was always performed simultaneously to ensure that the time and conditions under which the antibodies were bound to the cells was the same.

** PMA and ionomycin treatment resulted in death of less than 10% of thymocytes. This value represents IL-2R expression on the live cells only.

antibodies to CD4 and CD8 and second step reagents used to isolate the CD4⁺8⁺ thymocytes. Figure 2c illustrates the response of individual CD4⁺8⁺ sorted thymocytes to CD3. The level of [Ca²⁺]_i reached by the cells ranged from about 0.7–1.7 μ M. The smaller response relative to splenic T cells, could reflect the decreased amount of CD3 on the surface of CD4⁺8⁺ cells. As shown in Fig. 2d, 58% of the CD4⁺8⁺ thymocytes responded to antibody against CD3 by an increase of [Ca²⁺]_i above a threshold of 0.3 μ M. FACS analysis of the CD4⁺8⁺ cell preparation used for the experiment shown in Figs 2c and 2d indicated that 54% of the cells expressed CD3. The close agreement of the values for CD3 expression and for CD3-stimulated [Ca²⁺]_i mobilization indicates that the receptor complex on all CD3⁺, CD4⁺8⁺ cells can mediate calcium mobilization.

It has been established that protein kinase C (PKC), which is activated by diacylglycerol, the second messenger generated by phosphatidylinositol 4,5-bisphosphate hydrolysis, is also important in the activation of T cells. To determine if PKC is active in CD4⁺8⁺ thymocytes, we studied the effects of a phorbol ester, phorbol myristate acetate (PMA), known to activate PKC, on two cellular phenomena related to PKC activity. PMA induces rapid modulation of CD4 on T cells as a result of PKC-catalysed phosphorylation¹⁴. As shown in Fig. 3, PMA treatment greatly reduced cell surface expression of CD4, and of CD8 to a lesser extent, on CD4⁺8⁺ thymocytes. The modulation was specific, as expression of Lyt-1 was not affected by

PMA treatment (data not shown). It has also been established that PMA induces the phosphorylation of subunits of the CD3 complex¹⁵; we found that PMA induced phosphorylation of the ϵ and γ subunits of CD3 in the CD4⁺8⁺ thymoma line 1010, which is phenotypically and functionally equivalent to CD4⁺8⁺ thymocytes (data not shown). Taken together, these results indicate that PKC activity is present in CD4⁺8⁺ thymocytes and can be activated by PMA in the same manner as in functionally mature T cells.

As the components of the early steps of T cell activation appear to be in place in CD4⁺8⁺ thymocytes, it was of interest to determine the cellular response of CD4⁺8⁺ thymocytes to transmembrane signalling. As shown in Table 1, IL-2 receptor expression was induced on splenic T cells by antibody against CD3 and by the calcium ionophore ionomycin in the presence of PMA. But IL-2 receptor expression was not induced on CD4⁺8⁺ thymocytes by either of these stimuli. In the presence of irradiated macrophages, splenic T cells, or splenic T cells coated with antibodies against CD4 and CD8, responded to antibody to CD3 by IL-2 secretion and proliferation. Under similar conditions, CD4⁺8⁺ thymocytes failed to respond in either assay. The addition of 1 or 10 U/ml of recombinant IL-2 with antibody against CD3 did not lead to proliferation of the CD4⁺8⁺ thymocytes, suggesting that the lack of response is not simply due to the failure of these cells to secrete IL-2. As an additional test for functional response in the absence of

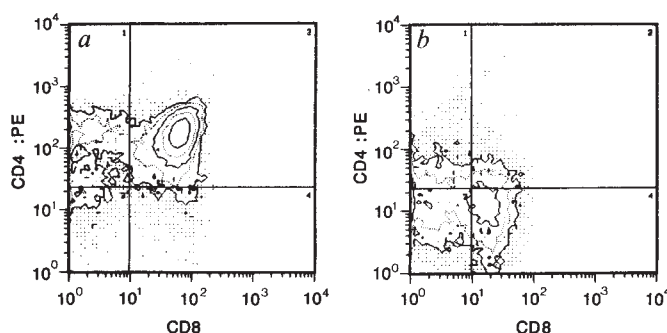


Fig. 3 Modulation of CD4 and CD8 on thymocytes by PMA. Thymocytes were incubated in medium alone (panel *a*) or in medium containing 10 ng ml^{-1} PMA for 6 h. The cells were then washed and stained for analysis of CD4 and CD8 expression as described in the legend to Fig. 1.

accessory stimuli, cells were treated with PMA in the presence of the calcium ionophore ionomycin. As shown in Table 1, IL-2 secretion and proliferation were observed in splenic T cells, but not in $\text{CD4}^+\text{CD8}^+$ thymocyte preparations. This result suggests that the failure of $\text{CD3}^+\text{CD4}^+\text{CD8}^+$ thymocytes to respond to CD3 antibodies by IL-2 secretion and proliferation cannot be attributed solely to the relatively small increase in $[\text{Ca}^{2+}]_i$ induced by antibody to CD3: the concentration of ionomycin used resulted in a sufficiently high mobilization of $[\text{Ca}^{2+}]_i$ to cause, with PMA, proliferation and lymphokine secretion in mature T cells. Together, these results demonstrate that CD3^+ , $\text{CD4}^+\text{CD8}^+$ thymocytes, although they express antigen receptors capable of mediating the early events associated with T-cell activation, cannot respond to appropriate stimuli by secretion of IL-2 or by proliferation.

The observation that a major subpopulation of murine $\text{CD4}^+\text{CD8}^+$ thymocytes expresses low quantities of CD3 is consistent with a previous analysis of human thymocytes¹⁶, and agrees with reports of expression of antigen receptor $V\beta$ gene products¹⁷ and CD3 (ref. 18) by murine thymocytes. The earliest thymocytes to appear in mice are $\text{CD4}^-\text{CD8}^-$, and lack the $\text{CD3-}\alpha/\beta$ receptor complex^{1,17,18}, although a subpopulation may express $\text{CD3-}\gamma/\delta$ receptor complex¹⁹. The next cells to appear express low levels of α/β receptors, detectable with an antibody to a determinant on the $V\beta 8$ family. The last cells to appear during ontogeny are mature cells expressing large amounts of antigen receptor¹⁷. Administration of antibody against receptor *in vivo*²⁰ or *in vitro*²¹ prevents the appearance of cells with high receptor expression, but has no effect on the appearance of cells with low receptor expression. More recently it has been shown that a $V\beta$ gene product which is associated with reactivity of the major histocompatibility complex class II protein, IE, is not detectable on mature cells in mice capable of expressing IE, but that the $V\beta$ gene is expressed at a normal frequency in thymocytes expressing low amounts of receptor^{22,23}. Together with our findings, this suggests that selection of antigen receptors could occur in single positive cells with high antigen receptor expression, rather than in the $\text{CD4}^+\text{CD8}^+$ cells with lower expression of antigen receptors. Further, these results indicate that an important maturation event could occur during any putative transition of $\text{CD4}^+\text{CD8}^+$ cells expressing low density receptor to high density mature cells. Our data suggest that the maturation event, if it occurs, is not related to the ability of the antigen receptor on the $\text{CD4}^+\text{CD8}^+$ cells to transmit an intracellular signal, but rather in how the cells perceive the signal. It is also possible that the inability of $\text{CD4}^+\text{CD8}^+$ cells to respond to antigen receptor-mediated stimuli by IL-2 secretion or by expression of IL-2 receptor renders the cells incapable of further maturation. An understanding of the events in thymocyte maturation will require further studies of the response of thymic subpopulations

to ligand-receptor interactions, and of the response of the cells to signals generated by these interactions.

This work was supported by grants from the NIH (J.P.A. and W.L.H.).

Received 21 May; accepted 14 September 1987.

1. Scollay, R., Bartlett, P. & Shortman, K. *Immun. Rev.* **82**, 79-103 (1984).
2. Adkins, B. *et al. A. Rev. Immun.* **5**, 325-365 (1987).
3. Fowlkes, B. J., Edison, L., Mathieson, B. J. & Chused, T. M. *J. exp. Med.* **162**, 802-822 (1985).
4. Reinherz, E. L. & Schlossman, S. F. *Cell* **19**, 821-827 (1980).
5. Allison, J. P., McIntyre, B. W. & Bloch, D. J. *Immun.* **129**, 2293-2300 (1982).
6. Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
7. Meuer, S. *et al. J. exp. Med.* **157**, 705-719 (1983).
8. Borst, J., Alexander, S., Elder, J. & Terhorst, C. *J. biol. Chem.* **258**, 5135-5143 (1983).
9. Allison, J. P. & Lanier, L. L. *Nature* **314**, 107-108 (1985).
10. Samelson, L. E., Harford, J. B. & Klausner, R. D. *Cell* **43**, 223-231 (1985).
11. Oettingen, H. C., Petteny, C. L., Maloy, W. L. & Terhorst, C. *Nature* **320**, 272-275 (1986).
12. Weiss, A. *et al. A. Rev. Immun.* **4**, 593-619 (1986).
13. Allison, J. P. *et al. The T Cell Receptor, UCLA Symposia on Molecular and Cellular Biology, New Series*, 73rd edn (eds Kappler, J. & Davis, M.) (Liss, New York, 1987).
14. Acres, R. B., Conlon, P. J., Mochizuki, D. Y. & Gallis, B. *J. biol. Chem.* **261**, 16210-16214 (1986).
15. Samelson, L. E., Patel, M. D., Weissman, A. M., Harford, J. B. & Klausner, R. D. *Cell* **46**, 1083-1090 (1986).
16. Lanier, L. L., Allison, J. P. & Phillips, J. H. *J. Immun.* **137**, 2501-2507 (1986).
17. Roehm, N. *et al. Cell* **38**, 577-584 (1984).
18. Bluestone, J. A., Pardoll, D., Sharrow, S. O. & Fowlkes, B. J. *Nature* **326**, 82-84 (1987).
19. Pardoll, D. *et al. Nature* **326**, 79-81 (1987).
20. McDuffie, M., Born, W., Marrack, P. & Kappler, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8728-8732 (1986).
21. Born, W. *et al. J. Immun.* **138**, 999-1008 (1987).
22. Kappler, J. W. *et al. Cell* **49**, 263-271 (1987).
23. Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
24. Mosmann, T. *J. immun. Meth.* **65**, 55-63 (1983).
25. Poenie, M., Alderton, J., Steinhardt, R. A. & Tsien, R. Y. *Science* **233**, 886-889 (1986).
26. Poenie, M., Tsien, R. Y. & Schmitt-Verhulst, A.-M. *EMBO J.* **6**, 2223-2232 (1987).

A new immunomodulating compound (AS-101) with potential therapeutic application

B. Sredni*, R. R. Caspi, A. Klein, Y. Kalechman, Y. Danziger, M. BenYa'akov, T. Tamari, F. Shalit & M. Albeck

Department of Life Sciences and Department of Chemistry, Bar-Ilan University, Ramat-Gan 52100, Israel

There has been interest in the potential of synthetic compounds to modify immune responses by imitation of cytokine action^{1,2}. Direct administration of interleukin 2 (IL-2) in conjunction with adoptive transfer of lymphokine activated killer cells has been used in the treatment of cancer^{3,4}, but there are toxic effects resulting from the high doses of IL-2 required^{3,5}. We have developed a new synthetic compound, ammonium trichloro(dioxoethylene-O,O'-tellurate) (AS-101), which has immunomodulating properties and minimal toxicity. The effects of AS-101 on the activation and function of immunocompetent cells have been assessed. We have found that AS-101 induces proliferation and IL-2 production by human lymphocytes *in vitro*, and enhances the production of IL-2 and colony-stimulating factor by mouse spleen cells. Splenocytes of BALB/c mice injected with AS-101 increased production of IL-2 and CSF *in vitro* in the presence of mitogen. Mononuclear cells of normal donors acquired responsiveness to recombinant IL-2 and bound monoclonal antibody to IL-2 receptor after incubation with AS-101. Splenocytes of mice treated *in vivo* with AS-101 expressed high levels of IL-2 receptor. The stimulation of lymphocytes by AS-101 apparently involves an increase in intracellular free calcium. AS-101 administered systemically to mice mediated antitumour effects which could be attributable to its immunomodulatory properties. In addition, AS-101 could directly enhance the ratio of OKT4 to

* To whom correspondence should be addressed at Department of Life Sciences, Bar-Ilan University, Ramat Gan, 52100, Israel.

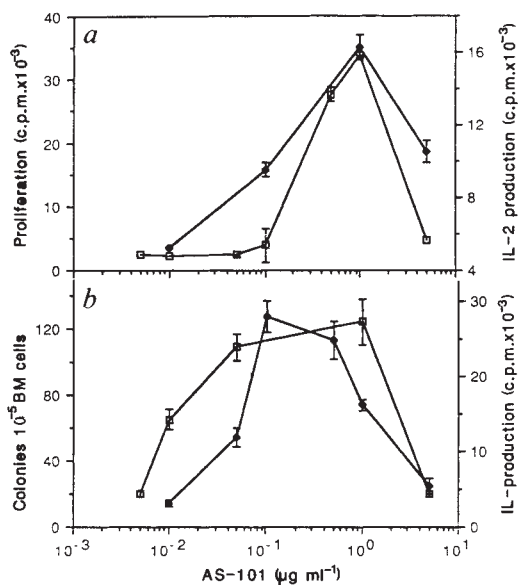


Fig. 1 Activation of human and mouse lymphoid cells by AS-101. *a*, Proliferation (open squares) and IL-2 production (filled squares) by human MNC incubated with AS-101 *in vitro*. *b*, *In vitro* induction of CSF (open squares) and IL-2 production (filled squares) by mouse spleen cells. IL-2 activity was measured by the growth of IL-2 dependent CTLL cells, as in *a*.

Methods. Human MNC were isolated by layering heparinized blood over a Ficol Hypaque gradient. The separated MNC were rinsed three times and brought to a concentration of $1\text{--}1.5 \times 10^6$ cells ml^{-1} in RPMI 1640 + 10% FCS with additions described³⁵. AS-101 was added to aliquots of the cell suspension and the cells seeded in triplicate 0.2 ml cultures in 96-well tissue culture trays. The cultures in which cell proliferation was to be assayed were incubated for 72 h and pulsed with $1 \mu\text{Ci}$ of ^3H -TdR per well during the last 14 h. Supernatants for assay of IL-2 activity were collected from parallel cultures at 48 h of incubation and assayed in triplicate on 1×10^4 cloned mouse IL-2-dependent CTLL cells in 0.2 ml cultures. The CTLL cultures were incubated for 40 h and pulsed with $1 \mu\text{Ci}$ ^3H -TdR per well during the last 14 h. Splenocytes from BALB/c mice (6–8 week old males) were brought to 10^7 cells ml^{-1} in supplemented RPMI 1640 and cultured with AS-101 as indicated. Supernatants were collected at 24 h and assayed for lymphokine activity. CSF activity was assayed as the number of colonies developing in semi-solid agar from murine BM cells⁹.

OKT8-positive cells in cultured mononuclear cells from AIDS (acquired immune deficiency syndrome) patients. These results indicate that AS-101 is potentially useful in the treatment of clinical conditions involving immunosuppression.

AS-101 is a low molecular weight organic tellurate compound soluble in organic solvents but only slightly soluble in water. Its development was based on a derivative of the antitumour agent dichloroethylenediamineplatinum (cisplatin)⁶. The elements, oxygen, sulphur and selenium have already been shown to have immunomodulatory activity^{1,7,8}. Selenium and tellurium were substituted for the platinum atom and screened for their ability to stimulate proliferation and lymphokine production by lymphoid cells, when it was found that the selenium derivatives were inactive in these assays and the tellurium compounds were active. The compound with the highest activity, ammonium trichloro (dioxoethylene-O,O') tellurate (coded AS-101), was selected for continued studies *in vivo* and *in vitro*.

AS-101 can stimulate normal lymphoid cells to proliferate and to produce lymphokines such as interleukin 2 (IL-2) and colony stimulating factor (CSF), which are regulators of lymphopoiesis and myelopoiesis^{9–11} (Fig. 1). Proliferation was measured directly by incorporation of tritiated thymidine (^3H -TdR) into DNA. IL-2 production was assayed by proliferation

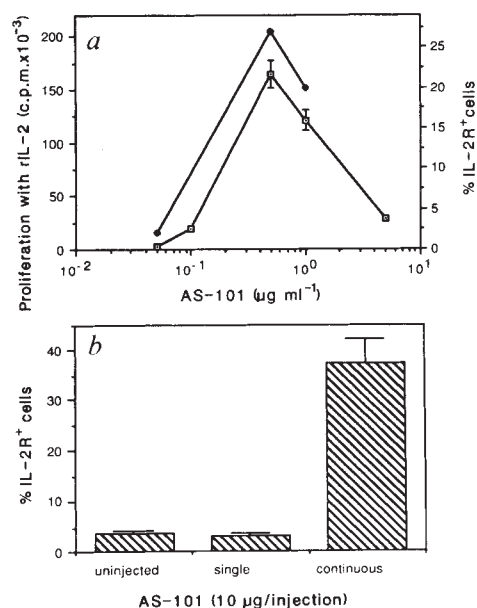


Fig. 2 IL-2 receptor induction by AS-101. *a*, *In vitro* induction in human MNC. MNC at 10^6 cells ml^{-1} were incubated with the indicated concentrations of AS-101 for 48 h and assayed for IL-2 receptor expression (filled squares) by immunocytofluorimetry (FACS, Becton Dickinson), as binding of anti-Tac monoclonal antibody (indirect method)³⁵. The anti-Tac was a gift from Dr. T. Waldmann. Expression of functional IL-2 receptors was assayed by responsiveness to rIL-2 of 10^6 MNC per ml incubated as 0.2 ml triplicate cultures for 24 h with the indicated concentrations of AS-101. The cells were washed in the plates and were resuspended in 0.2 ml of medium containing 20 U per ml of rIL-2 (Biogen) in triplicate 0.2 ml cultures (1 unit of IL-2 was defined as the concentration supporting half-maximal proliferation of CTLL cells). Proliferation was measured (open squares) after incubation of 48 h with rIL-2 by uptake of ^3H -TdR, added 14 h before harvest. *b*, Groups of BALB/c mice were injected intraperitoneally with 10 μg per animal of AS-101 given either as a single injection, or as continuous treatment, twice a week, for 2 months. Splenocytes were labelled by indirect immunofluorescence with a 1:1 mixture of monoclonal antibodies to the mouse IL-2 receptor, 2C7 and 7D4 (ref. 36), and analysed on a Cytofluorograf (Ortho) by comparison to splenocytes from untreated animals.

of an IL-2-dependent cloned murine CTLL line, and CSF production was estimated by the number of colonies developing in semi-solid agar from murine bone marrow cells (BM) in the presence of the supernatants from AS-101 stimulated cultures^{12,13}. Different concentrations of AS-101 in PBS were tested for induction of proliferation and IL-2 production in human mononuclear cells (MNC), and concentration of $1 \mu\text{g ml}^{-1}$ was found to be optimal for inducing both functions (Fig. 1*a*). AS-101 at $0.05\text{--}1 \mu\text{g ml}^{-1}$ induced optimal CSF production by mouse spleen cells, and optimal IL-2 production was found at $0.1\text{--}0.5 \mu\text{g ml}^{-1}$ AS-101 (Fig. 1*b*).

Proliferation of T cells in the presence of IL-2 depends on the simultaneous expression of functional IL-2 receptors^{20,21}. We therefore examined the induction of IL-2 receptors on MNC by their responsiveness to exogenous recombinant IL-2 (rIL-2) and by indirect immunofluorescence with the monoclonal anti-Tac antibody, which detects the IL-2 receptor antigen (Fig. 2*a*). MNC were pulsed with various concentrations of AS-101 and recultured with rIL-2. Cells preincubated with AS-101 at $0.5\text{--}1.0 \mu\text{g ml}^{-1}$ were induced to proliferate after the addition of rIL-2. This acquisition of responsiveness to IL-2 suggested that preincubation with AS-101 induced the expression of IL-2 receptor, an interpretation supported by the finding that some 25%

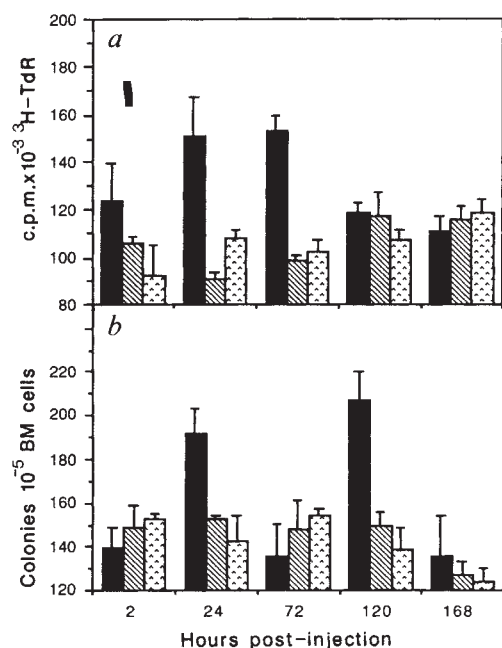


Fig. 3 Effect of *in vivo* treatment with AS-101 on subsequent *in vitro* lymphokine production by mouse spleen cells. 75 male BALB/c mice were injected once with 10 μ g AS-101 per animal. Fifteen animals were killed at each of the indicated times, and production of *a*, IL-2 and *b*, CSF by Concanavalin A-stimulated pooled splenocytes (2 μ g ml⁻¹, Difco) was assayed (filled bars) in comparison to the same number of uninjected controls (spotted bars) and to controls injected with the solvent only (cross-hatched bars). The assays were performed as described in the legend to Fig. 1.

of MNC from normal donors became anti-Tac positive under these conditions, as compared to only 2% of the cells incubated in medium alone (Fig. 2a).

IL-2 receptors are not expressed continuously by T cells, but only after appropriate immune stimulus²². For this reason, the dynamics of IL-2 receptor expression may determine the extent of the clonal expansion of T cells, which is a prerequisite for the induction of an *in vivo* immune response. The next step was therefore to test the effects of *in vivo* administration of AS-101. Groups of BALB/c mice were treated systemically with AS-101 at 10 μ g per animal, administered either as a single injection, or as a continuous treatment of twice-weekly injections for 2 months. AS-101 administered as a continuous treatment, but not as a single dose, was able to induce significant expression of IL-2 receptors *in vivo* (Fig. 2b). We have preliminary evidence for similar effects in human subjects. A single dose of 10 μ g of AS-101 injected into BALB/c mice caused a significant augmentation of IL-2 and CSF production in splenocytes explanted into culture and treated with mitogen (Fig. 3). IL-2 production was maximal at 1 and 3 days after injection (Fig. 3a), and two separate peaks for CSF activity were observed on days 1 and 5 (Fig. 3b). These two peaks represent the secretion of two different CSFs, the first being predominantly macrophage CSF and the second being mainly granulocyte/macrophage CSF (data not shown). The fact that *in vivo* treatment with AS-101 augments subsequent *in vitro* mitogenic responses suggests that AS-101 is an active biological response modifier that might increase both IL-2 and CSF production in response to a stimulus also *in vivo*.

As IL-2 and CSF have both separately been shown to promote tumouricidal activity²³⁻²⁵, we systemically administered AS-101 to model tumour-bearing C₃H/eb mice transplanted with a primary methylcholantrene-induced fibrosarcoma from the BALB/c strain. After 13 days, the tumour volume in untreated animals was 4-5-fold larger than in those treated with 10 μ g of

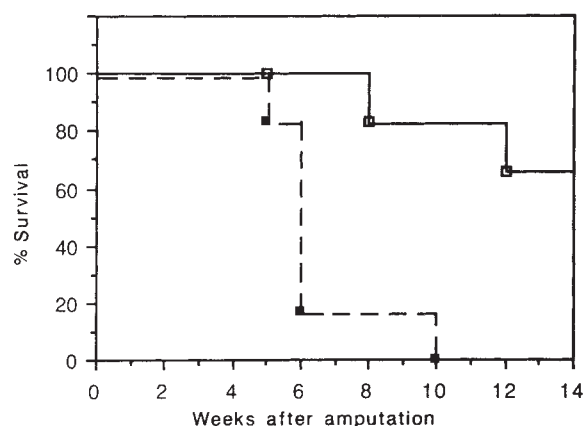


Fig. 4 Effect of AS-101 on survival of mice bearing Lewis lung carcinoma tumour metastases. C57Bl mice (groups of 12) were inoculated with 1×10^5 cells per mouse into one hind footpad. On day 12, when the tumours were easily palpable, the tumour-bearing footpads were amputated and twice-weekly intraperitoneal injections of 10 μ g AS-101 (open squares) or solvent only (filled squares) were started. Survival was followed for 3 months, and was significantly better in the treated group than in the control group ($p = 0.002$).

AS-101 administered twice weekly (data not shown). We also examined the effect of AS-101 on the survival of mice bearing metastases of the Lewis lung carcinoma, following removal of the primary tumour²⁶. The survival rate of the experimental animals as compared to that of the control group was greatly increased (Fig. 4). We conclude from our experiments that the antitumour activity of AS-101 could be attributed to its immunoenhancing properties; AS-101 might also have a direct antitumour effect because of its structural similarity to the antitumour agent cisplatin.

Toxicity tests showed that LD50 for AS-101 in rats was 5-10 mg kg⁻¹ for intravenous and 10 mg kg⁻¹ for intramuscular dosage, which is some 500-1,000 times higher than the immunologically effective dose. No acute toxicity of AS-101 was seen after intravenous administration of up to 2 mg kg⁻¹. In a subacute toxicity trial, rats given up to 2 mg kg⁻¹ of AS-101 three times weekly suffered no toxic effects; some renal toxicity was seen in animals receiving 4 mg kg⁻¹ six times weekly.

The immunomodulating and antitumour effects of AS-101, together with its minimal toxicity *in vivo*, make it a potentially useful drug. In AIDS patients, AS-101 was found to directly enhance the ratio of OKT4 to OKT8 positive cells *in vitro* in cultured MNC, as determined by immunocytofluorometry. After culture for 48 hours with 1 μ g ml⁻¹ of AS-101 the percentage of OKT4 positive cells rose from the mean basal level of 24 to 42, the percent of OKT8 positive cells remaining at the basal level of 39. MNC incubated in control medium remained at basal values. Current phase 1 clinical trials involve cancer and AIDS patients and bone-marrow transplant recipients, who show a general enhancement of immunological parameters, comparable to our results obtained *in vitro* and in animal models, without significant toxic effects at immunologically effective doses (unpublished data).

The mechanism of action of AS-101 in triggering lymphocyte stimulation resembles that of many other lymphocyte activators. Activation of lymphocytes depends on influx of Ca²⁺ through the cell membrane^{14,16}. Incubation of mouse splenocytes with AS-101 triggers a rapid 2.5-fold increase in intracellular Ca²⁺, as measured by the quin-2-acetoxymethyl ester assay^{16,27,28}. The specific calcium channel blocker nifedipine¹⁶ (30 μ M) or the divalent cation-chelating agent EGTA (2 Mm) (refs 14, 16) prevent the change in intracellular Ca²⁺. The biological effects of AS-101 on lymphocytes could involve the stimulation of Ca²⁺ influx, as nifedipine and EGTA both inhibit AS-101-induced IL-2 product by lymphoid cells (Table 1). Moreover, IL-2

Table 1 Effect of Ca^{2+} inhibitors on activation of lymphocytes by AS-101

Ca^{2+} inhibitor	AS-101 ($1 \mu\text{g ml}^{-1}$)	IL-2 titre (c.p.m. of CTLL)
Medium	+	28,195
EGTA (2 mM)	+	6,093
Nifedipine (30 μM)	+	14,540
None	—	918

Mouse splenocytes at 5×10^6 ml were preincubated in serum-free medium for 15 min with the Ca^{2+} inhibitor and then AS-101 was added for a further 1 h. The cells were washed and recultured in enriched RPMI+10% FCS for 24 h. Supernatants were collected and assayed for IL-2 titre on IL-2-dependent mouse CTLL cells. Cells recultured without AS-101 did not produce IL-2. In cultures where AS-101 was included in the reincubation, production of IL-2 by cells pretreated with the Ca^{2+} blockers was equal to that of non-pretreated control cells.

production induced by AS-101 is also suppressed by cyclosporin A, a fungal metabolite known to inhibit Ca^{2+} -dependent lymphocyte activation¹⁷⁻¹⁹ (data not shown).

In conclusion, AS-101 is a novel, minimally toxic immunomodulator, which induces IL-2 receptors and IL-2 and CSF production *in vitro* and *in vivo*. Thus, its direct or indirect cellular targets include T cells, macrophages and granulocytes. A CSF-mediated increase in macrophage/granulocytes could enhance the host response, to antigen administration, in tumour rejection, and in resistance to various infectious agents²⁴⁻²⁶. Potentiation of IL-2 production in conjunction with induction of IL-2 receptor expression could host defence, for example through NK activity, alloimmune response, tumour-specific response and LAK cells^{1,29,31}. Systemic administration of IL-2 is made difficult by its short half-life *in vivo* and by its toxicity at high doses^{3,5,32,33}. We suggest that immunomodulators, such as AS-101, that can induce endogenous IL-2 production *in vivo* would be more effective. Alternatively, AS-101 could be used to lower the doses of exogenous IL-2 because of its ability to induce the expression of IL-2 receptors.

Received 16 July; accepted 1 October 1987.

- Chirigos, M. A. & Talmadge, J. E. *Springer Seminar Immunopathol.* **8**, 327-346 (1985).
- Fenichel, R. L. & Chirigos, M. A. in *Immune Modulation Agents and their Mechanisms*, Vol. 25 (eds Fenichel, R. & Chirigos, M. A.) (Dekker, New York, 1985).
- Rosenberg, S. A. *et al. New Engl. J. Med.* **313**, 1485-1492 (1985).
- Mule, J. J., Shu, S., Schwarz, S. L. & Rosenberg, S. A. *Science* **225**, 1487-1489 (1984).
- Lotze, M. T. *et al. J. Immun.* **135**, 2865-2875 (1985).
- Rosenberg, S. A. *Cancer* **55**, 2303-2316 (1985).
- Brown, R. L., Griffith, R. L., Ruscetti, F. W. & Rabin, H. *Cell. Immun.* **92**, 14-21 (1985).
- Goodman, M. G. & Weigle, W. O. *J. Immun.* **126**, 20-26 (1981).
- Farrar, J. J. *et al. Immunol. Rev.* **63**, 129-166 (1982).
- Oppenheim, J. J. & Rosenstreich, D. L. *Prog. Allergy* **20**, 65-194 (1976).
- Apte, R. N., Heller, E., Hertogs, C. F. & Pluznik, D. H. in *Experimental Hematology Today* (eds Baum, S. J. & Ledney, D. G.) 213-220 (Springer, New York, 1978).
- Pluznik, D. H. & Sachs, L. *J. cell comp. Physiol.* **66**, 319-329 (1965).
- Bradley, T. R. & Metcalf, D. J. *Aust. J. exp. Biol. med. Sci.* **44**, 287-299 (1966).
- Mills, G. B., Cheung, R. K., Grinstein, S. & Gelfand, E. W. *J. Immun.* **134**, 1640-1643 (1985).
- Truneh, A., Albert, F., Golstein, P. & Schmitt-Verhulst, A.-M. *Nature* **313**, 318-320 (1985).
- Gelfand, E. W., Cheung, R. K., Grinstein, S. & Mills, G. B. *Eur. J. Immun.* **16**, 907-912 (1986).
- Manger, B., Hardy, K. J., Weiss, A. & Stobo, J. D. *J. clin. Invest.* **77**, 1501-1506 (1986).
- Kay, J. E. & Benzie, C. R. *Immunology* **57**, 195-199 (1986).
- Gelfand, E. W., Cheung, R. K. & Mills, G. B. *J. Immun.* **138**, 1115-1120 (1987).
- Robb, R. J., Munk, A. & Smith, K. A. *J. exp. Med.* **154**, 1455-1474 (1981).
- Smith, K. A., Wang, H. M. & Cantrel, D. A. in *Progress in Immunology* (eds Yamamura, Y. Y. & Tada, T.) 269 (Academic, New York, 1983).
- Cantrel, D. A. & Smith, K. A. *J. exp. Med.* **158**, 1895-1911 (1983).
- Rosenberg, S. A., Mule, J. J., Spiess, P. J., Reichert, C. M. & Schwarz, S. L. *J. exp. Med.* **161**, 1169-1188 (1985).
- Grabstein, K. H. *et al. Science* **232**, 506-508 (1986).
- Sato, M., Ichimura, O., Mitsuno, T., Koyima, E. & Osawa, T. *J. pharm. Dyn.* **5**, 818-828 (1982).
- Klein, A. & Shoham, J. *Cancer Res.* **41**, 3217-3221 (1981).
- Gelfand, E. W., Cheung, R. K. & Grinstein, S. *J. cell Physiol.* **121**, 533-539 (1984).
- Tsein, R. Y., Pozzan, T. & Rink, T. J. *J. Cell Biol.* **94**, 325-334 (1982).
- Rosenberg, S. A., Spiess, P. J. & Schwarz, S. *Transplantation* **35**, 631-634 (1983).
- Chang, A. E., Hyatt, C. L. & Rosenberg, S. A. *J. Biol. Resp. Mod.* **3**, 561-572 (1984).
- Rosenberg, S. A., Spiess, P. J. & Lafreniere, R. *Science* **233**, 1318-1321 (1986).
- Bindon, C. *et al. Br. J. Cancer* **47**, 123-133 (1983).
- Lotze, M. T., Frana, W., Sharrow, S. O., Robb, R. J. & Rosenberg, S. A. *J. Immun.* **134**, 157-166 (1985).
- Sredni, B., Friedman, M. M., Bland, C. E. & Metcalfe, D. D. *J. Immun.* **131**, 915-922 (1983).
- Tsuda, M., Uchiyama, T., Takatsuki, K., Uchino, H. & Yodoi, J. *J. Immun.* **129**, 592-595 (1982).
- Lillehoj, H. S., Malek, T. R. & Shevach, E. M. *J. Immun.* **133**, 244-250 (1984).

Differential expression of calmodulin-binding proteins in B, T lymphocytes and thymocytes

Randall L. Kincaid*†, Hajime Takayama†, Melvin L. Billingsley‡ & Michail V. Sitkovsky†

* Laboratory of Cellular Metabolism, National Heart, Lung, and Blood Institute and †Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA
‡ Milton S. Hershey Medical Center, Pennsylvania State University, Hershey, Pennsylvania 17033, USA

Changes in intracellular free Ca^{2+} are involved in the transmembrane signalling of different cells, including lymphocytes¹⁻³. Since calmodulin (CaM) is a primary receptor for Ca^{2+} (ref. 4), it may mediate the activation of crucial enzymes after antigen-induced increases in cytosolic Ca^{2+} . Using a biotinylated-CaM (Bio-CaM) detection procedure⁵ to identify such proteins, we found that a peptide of relative molecular mass 59,000 (59K) was the predominant soluble CaM-binding protein (CaM-BP) in T cells and B lymphocytes from murine spleen; immunoblotting experiments identified it as a subunit of the CaM-dependent phosphatase, 'calcineurin' (CN)⁶. Smaller amounts of larger CaM-BPs, thought to be cytoskeletal-binding proteins, were also detected. CaM-BPs were expressed differentially, with B lymphocytes having four times more of the CN-like protein than T lymphocytes, while in thymocytes, a 65K polypeptide was the major CaM-BP. However, limited proteolysis analysis suggested that this thymus-specific peptide may be a precursor of CN. These data suggest that Ca^{2+} -stimulated protein dephosphorylation may be an important and highly regulated function in lymphoid cells.

Initial experiments indicated very low amounts of soluble CaM-BPs in isolated murine splenocytes, necessitating their enrichment prior to analysis. Cell extracts were depleted of endogenous CaM by anion-exchange chromatography⁷ and CaM-BPs were concentrated by CaM-Sepharose affinity chromatography (see Fig. 1 legend). Following SDS-polyacrylamide gel electrophoresis of the CaM-Sepharose eluate, polypeptides of 59K and 43K were detected (Fig. 1a); of these only the 59K band reacted with the biotinylated derivative of CaM (Bio-CaM)⁵ on Western blots (Fig. 1b). The unreactive 43K peptide probably represented nonspecific retention of cytosolic protein (for example, actin) on the affinity gel, although it may be a labile CaM-BP, the binding activity of which was damaged by denaturation during gel electrophoresis. CaM-dependent cyclic nucleotide phosphodiesterase^{8,9}, CaM-dependent protein kinase^{10,11}, and the CaM-dependent protein phosphatase calcineurin (CN)^{12,13} all have subunits of ~60K. Therefore, it was important to distinguish whether one or more was present in this molecular mass region. Immunoblotting experiments established that only the antibody to CN reacted with the 59K polypeptide, while no reaction was seen with antibodies against phosphodiesterase (Fig. 1c) or protein kinase (P. Kelly, R.K. and M.S., data not shown). Since the limit for immunodetection of phosphodiesterase and kinase was 5 ng, and the amount of CN-immunoreactive peptide was 50-60 ng (based on an internal standard of CN), it appeared that the CN-like protein comprised at least 90% of Bio-CaM-reactive peptide with this size.

These results indicate that the major CaM-binding peptide in murine lymphocytes is the catalytic subunit of CaM-dependent protein phosphatase. In some preparations, small amounts of higher relative molecular mass peptides (220 and 150K), were observed (see Fig. 2). The sizes of these peptides corresponded

† Present address: Section on Immunology, National Institute of Alcohol Abuse and Alcoholism, 12501 Washington Avenue, Rockville, Maryland 20852, USA.

Fig. 1 Analysis of CaM-BPs from isolated murine spleen cells by SDS-PAGE and Western blot procedures. **a**, Coomassie blue staining of SDS-gel (13% acrylamide). Lane 1, protein standards (Pharmacia); lane 2, 10 μ g of material applied to CaM-Sepharose; lane 3, CaM-Sepharose eluate, 0.3 μ g. **b**, Bio-CaM reactivity of Western blot. Lanes 1–3, amido black staining: lane 1, pre-stained protein standards (BRL); lane 2, bovine brain CN, 0.8 μ g; lane 3, CaM-Sepharose eluate (0.2 μ g). Lanes 4–6, Bio-CaM reactivity; lanes 4, 5, the same fractions as in lanes 2, 3 (Fig. 1b); lane 6, the same fraction as lane 5 but incubated in the presence of 1 mM EGTA (to demonstrate that Bio-CaM binding was Ca^{2+} -specific). **c**, Immunoreactivity toward anti-CN and anti-phosphodiesterase (PDE) antibodies. Lanes 2, 4, 0.3 μ g of CaM-Sepharose eluate. Lanes 1, 3, pre-stained standards containing 50 ng of CN or PDE, respectively. Lanes 1, 2; CN immunoreactivity, lane 3, 4; PDE immunoreactivity.

Methods. Spleens from 100 BALB/c mice were depleted of red blood cells by hypotonic shock and purified by Ficol-Hypaque centrifugation¹⁹. Cells were homogenized in 40 mM Tris-HCl pH 7.4, 1 mM ZnCl_2 , 1 mM PMSF, 10 $\mu\text{g ml}^{-1}$ antipain, 10 $\mu\text{g ml}^{-1}$ leupeptin and the homogenate centrifuged at 12,000g for 10 min; the soluble fraction was dialysed against 25 mM Tris-HCl, pH 8.0 containing 0.2 M NaCl and 1 mM EGTA. QAE-Sephadex A-25 (Pharmacia) was added (~3 ml gel) and the suspension was mixed for 2–4 h at 4 °C; the fraction not retained on QAE was collected (12 ml, 6.5 mg). This CaM-depleted cytosol was adjusted to 2 mM free Ca^{2+} , and mixed with CaM-Sepharose (0.5 ml gel) for 1–2 h. After washing with Ca^{2+} -containing buffer which included detergent the affinity gel was eluted with EGTA⁹. The eluate (1.5 ml, 1.2 μg) was concentrated by TCA precipitation and solubilized in SDS-containing buffer. Gels were electroblotted and nitrocellulose sections incubated with Bio-CaM⁵ or affinity-purified antibodies to PDE and CN²⁴. After washing, bound Bio-CaM was detected with avidin-horseradish peroxidase (Vector Labs). PDE and CN antibodies were detected similarly, after prior incubation with biotinylated second antibody.

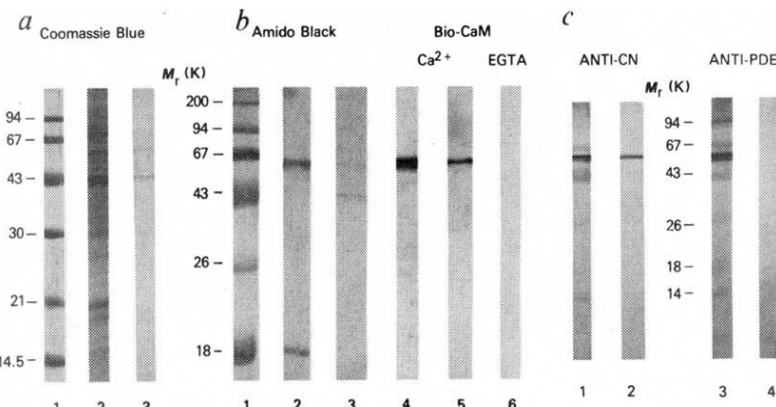
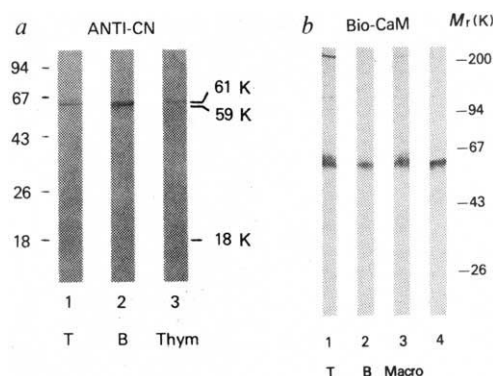


Fig. 2 CaM-binding proteins and anti-CN immunoreactivity in murine T, B lymphocytes, thymocytes and macrophages. **a**, Anti-CN immunoreactivity in T, B lymphocytes and thymocytes. Lane 1, T lymphocytes; lane 2, B lymphocytes; lane 3, thymocytes. Equivalent cell numbers of T, B lymphocytes and thymocytes were used. Note that the immunoreactive peptide from thymocytes (61K) was somewhat larger than that from B or T cells; bovine brain CN (in the same experiment) migrated with an apparent M_r of 61K. **b**, CaM-binding proteins in murine T, B lymphocytes and macrophages. Lane 1, T lymphocytes; lane 2, B-lymphocytes; lane 3, macrophages; lane 4, pre-stained protein standards containing 50 ng purified CN. Equivalent numbers of T cells and macrophages were analysed, however, only 20% as many B lymphocytes were used.

Methods. Spleen cells and thymocytes were prepared as described in Fig. 1 legend and ref. 19. Spleen cells from 100 mice (6.5×10^9 cells) were incubated in plastic culture dishes for 1 h at 4 °C and non-adherent cells were harvested and suspended in 200 ml of RPMI-1640/20 mM HEPES pH 7.2; adherent cells (2.2×10^9 cells) were used as macrophages. T and B lymphocytes were purified by plating them onto plastic culture dishes coated with goat anti-mouse IgM (kindly provided by Dr Wayne Tsang) for 1 h at 4 °C. The non-adherent T cell population was treated for 45 min (37 °C) with anti-Ia monoclonal antibody (M5/114 hybridoma²⁰ supernatant) and guinea pig complement to lyse contaminating B cells. Spleen cells that adhered to anti-IgM coated plates were used as B cells. As precautions were taken to remove contaminating B cells from T cell preparations, the Bio-CaM and anti-CN reactivity observed in T cells could not be attributed to contamination by B cells. However, because complete removal of B cells from the adherent (macrophage) population is difficult, some CaM-BPs in macrophages may arise from B cells. Procedures for identification of CaM-BPs and CN immunodetection are described in Fig. 1 legend, except that for panel b, detection of Bio-CaM was carried out using avidin-alkaline phosphatase (Vector).



to those of two known CaM-BPs, spectrin¹⁴ and caldesmon^{15,16}, which are thought to associate with membranes and actin in a Ca^{2+} -dependent fashion. Taken together, these findings suggest an important role for phosphoprotein dephosphorylation (and, perhaps, cytoskeletal regulation) in Ca^{2+} and CaM-dependent processes in lymphocytes.

The average yield of 59K peptide was ~600–800 ng per $\sim 10^{10}$ lymphocytes, making extensive characterization of the protein difficult. However, using antibody to CN, we compared the relative amounts of immunoreactive protein in purified preparations of T, B lymphocytes, macrophages and thymocytes. Interestingly, B lymphocytes contained about four times more of this protein than T cells or thymocytes (Fig. 2a). In this experiment, an immunoreactive 18K peptide was observed in the B cell CaM-BP fraction (lane 2), further confirming its identity with CN, which also contains an 18K Ca^{2+} -binding subunit⁶. The difference in content of immunoreactive 59K protein in T and B cells was also reflected in the amount of Bio-CaM reactivity

(Fig. 2b). When five times fewer cell equivalents of B lymphocytes were compared, similar amounts of the major CaM-BP were observed for B and T cells; the amount of this CaM-BP in macrophages was also less than in B lymphocytes, but similar to T cells (lane 3). These findings suggest significant differences in Ca^{2+} -dependent phosphoprotein metabolism between T and B lymphocytes. It is known that specific proteins are differentially phosphorylated in T and B lymphocytes in response to mitogenic stimulation¹⁷ and that distinct protein kinases are expressed in these cells.¹⁸ The relationship of such differences to the amount of CN in T, B cells is not clear; however, such differential expression of CaM-dependent phosphatase may reflect differences in the biochemical pathways for Ca^{2+} -dependent activation of T, B lymphocytes.

Surprisingly, the 59K CN-immunoreactive peptide was not the predominant CaM-BP in thymocytes. Instead, the major CaM-BP was a 65K peptide that did not cross-react with anti-CN antibody (Fig. 3a). Other CaM-BPs in thymocytes included CN,

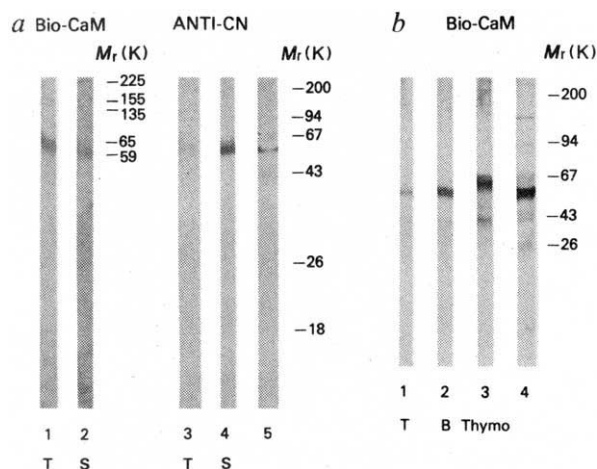


Fig. 3 Comparison of CN immunoreactivity and CaM-BPs from murine lymphocytes and thymocytes. *a*, Bio-CaM and anti-CN reactivity in the CaM-Sepharose eluates from murine thymus and spleen. Identical samples of CaM-Sepharose eluates were electroblotted onto nitrocellulose and analysed for Bio-CaM and anti-CN reactivity. Lanes 1, 3, thymocytes; lanes 2, 4, spleen cells; lane 5, pre-stained protein standards containing 40 ng calcineurin. Lanes 1, 2-Bio-CaM reactivity; lanes 3, 4, 5-immunoreactivity with anti-CN antibody. *b*, CaM-binding proteins in T, B lymphocytes and thymocytes. Identical samples used in Fig. 2*a* (for detection of CN immunoreactivity), were analysed for Bio-CaM reactivity. Lane 1, T lymphocytes; lane 2, B lymphocytes; lane 3, thymocytes; lane 4, pre-stained protein standards, containing 50 ng of CN.

Methods. All procedures as described in Fig. 1 legend, except that the detecting reagents for Bio-CaM and anti-CN antibody were avidin-alkaline phosphatase and anti-goat antibody conjugated to alkaline phosphatase (Kirkegaard-Perry), respectively.

detected by immunoblotting experiments, and the previously-mentioned cytoskeletal CaM-binding proteins (Fig. 3*a, b*). In several thymocyte preparations, the 65K peptide appeared to be partially degraded into peptides of 63 and 61K (Fig. 4, lane 2). These peptides showed significant CN immunoreactivity, albeit less than that observed with a comparable amount of the spleen peptide of 59K (Fig. 4, lanes 3, 4). These observations suggested that the 65K 'thymus-specific' CaM-BP might represent a larger form of CN that is not immunoreactive, and that degradation produced slightly smaller forms of increased immunoreactivity. When CaM-BPs from thymus and spleen were partially digested with staphylococcal V-8 protease, virtually identical Bio-CaM binding peptides were observed in these fractions and in isolated T cells (Fig. 4*b*). These findings are consistent with a relationship wherein the thymus-specific protein represents a precursor form of CN; such a modified species may be 'processed' into smaller forms of this enzyme

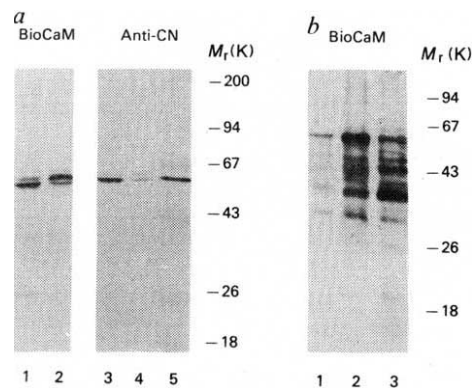


Fig. 4 CaM-binding peptides from murine spleen and thymus before and after limited proteolysis with staphylococcal V-8 protease. *a*, Bio-CaM and anti-calcieneurin reactivity in CaM-Sepharose eluates from murine thymus and spleen. Samples, from a different preparation, were treated as described in the legend to Fig. 3*a*. Lane 1, 3 spleenocytes; lane 2, 4 thymocytes; lane 5, pre-stained protein standards containing 50 ng of CN. *b*, Bio-CaM reactivity after treatment of samples with staphylococcal V-8 protease. The CaM-Sepharose eluates, used for *a* and from T lymphocytes (Fig. 3*b* above), were incubated at 37° for 30 min with *S. aureus* V-8 protease (1:20, w/w of protein substrate), after which portions were electrophoresed on an acrylamide gel (15%) and electroblotted; the nitrocellulose was then incubated to detect Bio-CaM binding proteins as outlined previously. Lane 1, T lymphocytes; lane 2, spleenocytes; lane 3, thymocytes.

during differentiation of thymocytes into functionally mature T lymphocytes.

Although originally thought to occur primarily in nervous tissue⁶, CN has now been demonstrated in many other tissues²¹, including cells of lymphoid origin, as shown by Chantler²². Although CN is mainly a soluble protein, immunochemical studies of brain^{23,24} and mesenteric lymph node cells²² suggest its association with membrane structures as well. The latter study made the provocative observation that a somewhat larger immunoreactive form (68K) could be extracted from lymphocyte membranes prepared with Ca²⁺, but not from those prepared with EGTA. Thus, this protein phosphatase may associate reversibly with cellular membranes, acting on both soluble and membrane-bound phosphoprotein substrates²⁴. Since phosphorylation of specific proteins and increases in intracellular Ca²⁺ are known to accompany lymphocyte activation, the present finding that CN is the major CaM-BP implies that protein dephosphorylation may play an important role in such events.

This work was supported in part, by a research grant from the International Life Sciences Institute Research Foundation to M.L.B.

Received 24 November 1986; accepted 28 September 1987.

- Allwood, G., Asherson, G. L., Davey, M. J. & Goodford, P. J. *Immunology* **21**, 509-516 (1971).
- Freedman, M. H. & Raff, M. C. *Nature* **255**, 378-382 (1975).
- Tsien, R. Y., Pozzan, T. & Rink, T. J. *Nature* **295**, 68-71 (1982).
- Means, A. R. & Dedman, J. R. *Nature* **285**, 73-78 (1980).
- Billingsley, M. L., Pennypacker, K. R., Hoover, C. G., Brigati, D. J. & Kincaid, R. L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7585-7589 (1985).
- Klee, C. B., Crouch, T. H. & Krinks, M. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6270-6273 (1979).
- Kincaid, R. L. & Vaughan, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4903-4907 (1979).
- Sharma, R. K., Wang, T. H., Wirch, E. & Wang, J. H. *J. biol. Chem.* **255**, 5916-5923 (1980).
- Kincaid, R. L. *et al. J. biol. Chem.* **259**, 5158-5166 (1984).
- Bennett, M. K., Erond, N. E. & Kennedy, M. B. *J. biol. Chem.* **258**, 12735-12744 (1983).
- Goldenring, J. R., Gonzalez, B., McGuire, J. S. Jr & DeLorenzo, R. J. *J. biol. Chem.* **258**, 12632-12640 (1983).
- Klee, C. B. & Krinks, M. H. *Biochemistry* **17**, 120-126 (1978).

- Wallace, R. W., Lynch, T. J., Tallant, E. A. & Cheung, W. Y. *J. biol. Chem.* **254**, 377-382 (1979).
- Sobue, K., Muramoto, V., Fujita, M. & Kakiuchi, S. *Biochem. biophys. Res. Commun.* **100**, 1063-1070 (1981).
- Sobue, K., Muramoto, Y., Fujita, M. & Kakiuchi, S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5652-5655 (1981).
- Bretscher, A. *J. biol. Chem.* **259**, 12873-12880 (1984).
- Dasch, J. R. & Stavitsky, A. B. *Molec. Immun.* **22**, 379-389 (1985).
- Harrison, M. L., Low, P. S. & Geahlen, R. L. *J. biol. Chem.* **259**, 9348-9350 (1984).
- Mishel, B. B. & Shiigi, S. M. (eds) in *Selected Methods in Cellular Immunology* (Freeman, San Francisco, 1980).
- Bhattacharya, A., Dorf, M. E. & Springer, T. A. *J. Immun.* **127**, 2488-2495 (1981).
- Pallen, C. J. & Wang, J. H. *Archs Biochem. Biophys.* **237**, 281-291 (1985).
- Chantler, P. D. *J. Cell Biol.* **101**, 207-216 (1985).
- Wood, G. C., Wallace, R. W., Whitaker, G. N. & Cheung, W. Y. *J. Cell Biol.* **84**, 66-76 (1980).
- Kincaid, R. L., Balaban, C. D. & Billingsley, M. L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1118-1122 (1987).

Both immature and mature T cells mobilize Ca^{2+} in response to antigen receptor crosslinking

Terri Helman Finkel, Marcia McDuffie,
John W. Kappler, Philippa Marrack & John C. Cambier

Department of Immunology, National Jewish Center for Immunology and Respiratory Medicine, 1400 Jackson Street, Denver, Colorado 80206, USA

T cells develop from prothymocytes which express no detectable antigen receptors to immature thymocytes with few receptors, eventually becoming mature thymocytes and peripheral T cells with 20,000–40,000 receptors per cell¹. Recent studies^{2–8} suggest that immature thymocytes are immunologically unresponsive. We have suggested^{2,3} that an early step in signal transduction following engagement of the T cell receptor might differ in immature and mature T cells. Here we examine anti-receptor antibody mediated induction of calcium mobilization in immature and mature T cells. Results indicate that antigen receptors on both immature and mature receptor-positive T cells transduce signals via calcium mobilization. Significant differences were observed, however, between these populations in the magnitude of influx of extracellular Ca^{2+} following binding of antireceptor antibody. Specifically immature cells show a much reduced Ca^{2+} influx response compared to mature cells which could result from a low Ca^{2+} channel⁹ frequency in the plasma membranes of immature T cells, or from less efficient activation of existing channels.

Activation of mature T lymphocytes is initiated through interaction of ligand or antibody with the T cell antigen receptor complex (TcR or T3) (refs 10–12). This interaction stimulates

Ca^{2+} mobilization and proliferation in the presence of IL-2 (refs 13–15). We report here the induction of Ca^{2+} mobilization in mature and immature T cells by crosslinking of antigen receptors in the absence of additional stimuli. We have used the monoclonal anti-T cell receptor antibody, F23.1 (F23) (ref. 16), which reacts with the three products of the V β 8 gene family (V β 8.1, V β 8.2, and V β 8.3). This antibody specifically stains 30–40% of peripheral T cells and 15% of thymocytes of strain C3H/HeJ. As shown in Fig. 1, the percentage of cells in C3H/HeJ thymus and lymph node which respond to crosslinking of TcR by F23, with an increase in intracellular free calcium ($[\text{Ca}^{2+}]_i$), corresponds with the percentage of cells identified by immunofluorescence analysis as F23-binding. Addition of goat anti-mouse immunoglobulin (GAMIg), to F23-coated peripheral T cells of strain C3H/HeJ, of which $32.1 \pm 0.7\%$ ($n=5$) bind F23 by fluorescent staining (Fig. 1b), or thymocytes, of which $14.6 \pm 1.7\%$ ($n=7$) are F23⁺, resulted in an increase in $[\text{Ca}^{2+}]_i$ by $28.3 \pm 1.0\%$ ($n=12$) and $12.3 \pm 0.6\%$ ($n=25$), of cells respectively (Fig. 1a). Thus, crosslinking of the antibody-TcR complex results in an increase in $[\text{Ca}^{2+}]_i$ in virtually all F23⁺ cells from thymus and periphery.

As illustrated in Fig. 1b, fluorescent staining has identified two populations of antigen receptor-bearing thymocytes¹⁷. A small population (3–5%) of cells bear high numbers of F23⁺ receptors and are therefore considered mature¹⁷. A percentage of antigen receptor-bearing cells (12–15% of total cells) express F23⁺ receptors in low numbers. These are the immature thymocytes (ref. 17, and unpublished observations for C3H/HeJ). The observation that the 15% of C3H thymocytes which express F23⁺ receptors mobilize Ca^{2+} in response to crosslinking of the anti-TcR antibody, F23, indicates that the TcR on both mature (3%) and immature (12%) receptor-positive cells is coupled to Ca^{2+} mobilization.

Two approaches were taken to corroborate this evidence.

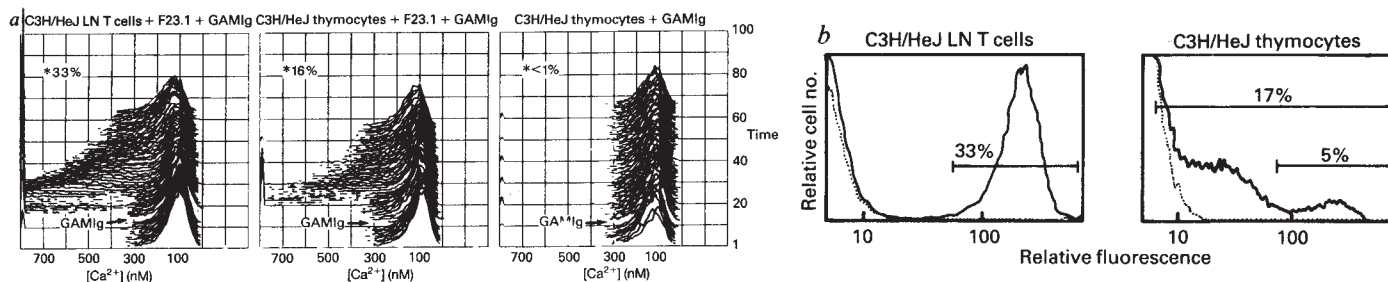


Fig. 1 Increase in intracellular free Ca^{2+} levels ($[\text{Ca}^{2+}]_i$) induced by crosslinking of anti-T cell receptor (TcR) antibody, F23, compared with the frequency of receptor-positive cells by fluorescent staining. **a**, Isometric display illustrating calcium mobilization in response to addition of goat anti-mouse immunoglobulin (GAMIg) crosslinking reagent to cells bound with anti-TcR antibody, F23. Flow cytometric analysis was begun before addition of GAMIg to obtain an unstimulated baseline measurement. The cell flow was halted for addition of GAMIg in excess, at the time indicated by arrows, resulting in a gap in the profiles. Asterisks indicate % of cells which responded with increased $[\text{Ca}^{2+}]_i$. In general, lymph node T cells responded maximally above a threshold value of the indo-1 ratio two standard deviations above the mean resting ratio. In general, thymocytes responded maximally above a threshold ratio value one standard deviation above the mean resting ratio. These data were confirmed by curve subtraction of baseline from responding populations using data stored in 'list-mode'. Data shown are representative of 10 experiments; 7 channels represent 1 min on the timescale. **b**, Immunofluorescence staining of cells positive for the T cell receptor for antigen. Cells positive for V β 8 TcR are denoted by the solid line (—); cells negative for this TcR (background staining) are denoted by the dotted line (···). The % shown represent the proportion of F23⁺ cells falling in the integrated area beneath the bar.

Methods. **a**, Isometric display. Nylon wool purified lymph node cells or unfractionated thymocytes from C3H/HeJ adult mice were loaded with indo-1, as described²². Cells were incubated for 30 min at 37 °C with or without the primary antibody, F23 ($10 \mu\text{g ml}^{-1}$), washed three times, suspended in balanced salt solution (2×10^6 cells ml^{-1} per sample) for Ca^{2+} measurements, and equilibrated at 37 °C for 15 min. $[\text{Ca}^{2+}]_i$ was assayed cytofluorometrically²³. GAMIg ($100 \mu\text{g ml}^{-1}$, Jackson) was added as indicated. Calibration of $[\text{Ca}^{2+}]_i$ was as described²³. Analyses were performed using an Ortho Cytofluorograph 50H with the model 2150 computer. Data were collected using the 2150 computer 'time mode' in which indo-1 violet/blue fluorescence ratio (y axis) was displayed as a function of elapsed time (x axis). Programs²³ were used for later analysis of stored data to calculate the mean indo-1 fluorescence ratio against time, and percentage of responding cells against time. **b**, Immunofluorescence staining. Stained samples were prepared as above, then incubated with fluorescein-conjugated GAMIg (—) or with fluorescein-conjugated GAMIg alone (···), and analysed by flow cytometry. The anti-receptor antibody concentration dependence for activation of this response paralleled that for staining, with the optimal concentration being $1 \mu\text{g per ml per } 10^6$ cells (data not shown). Addition of GAMIg to cells not pretreated with anti-receptor antibody induced no change in $[\text{Ca}^{2+}]_i$ (Fig. 1a). Crosslinking of the TcR using GAMIg was required for the Ca^{2+} response as addition of F23 alone to cells resulted in no change in $[\text{Ca}^{2+}]_i$ (data not shown).

Fig. 2 Comparison of Ca^{2+} mobilization in immature and mature thymocytes in response to engagement of the T cell antigen receptor. *a* and *b*, Change in $[\text{Ca}^{2+}]_i$ with time upon receptor crosslinking for (a) newborn and (b) adult thymocytes. Asterisks indicate percentages of cells which responded with increased $[\text{Ca}^{2+}]_i$ as described in the legend to Fig. 1. 7 channels represent 1 min on the timescale. *c*, Percentages of total cells which responded with increased $[\text{Ca}^{2+}]_i$, newborn (open squares) and adult (filled squares) thymocytes, determined as described in the legend to Fig. 1. Immunofluorescence staining data for newborn (*d*) and adult thymocytes (*e*). The %s represent the proportion of F23^+ cells falling in the integrated area beneath the bar. The representative data shown in Fig. 2 are highly reproducible. A compilation of this data shows that at the time of peak response (90–150 s), $11.4 \pm 0.6\%$ ($n=6$) of newborn thymocytes and $12.3 \pm 0.6\%$ ($n=25$) of adult thymocytes respond to crosslinking of TcR with an increase in $[\text{Ca}^{2+}]_i$. In experiments done in parallel, $12.1 \pm 0.4\%$ ($n=6$) of newborn thymocytes and $14.6 \pm 1.7\%$ ($n=7$) of adult thymocytes bound F23 by fluorescent staining. Experiments done with mice at days 19 and 18 of gestation, with a high density population of 0.5% and $<0.5\%$ respectively, gave similar results (data not shown).

Methods. Panels *a*, *b* and *c*, Thymocytes from C3H/HeJ newborn (days 20–21 of gestation) and adult mice were loaded with indo-1 and incubated with F23. GAMlg was added at time 0 (Fig. 2c) or as indicated by arrows. Panels *d* and *e* Cells were stained as described in the legend to Fig. 1.

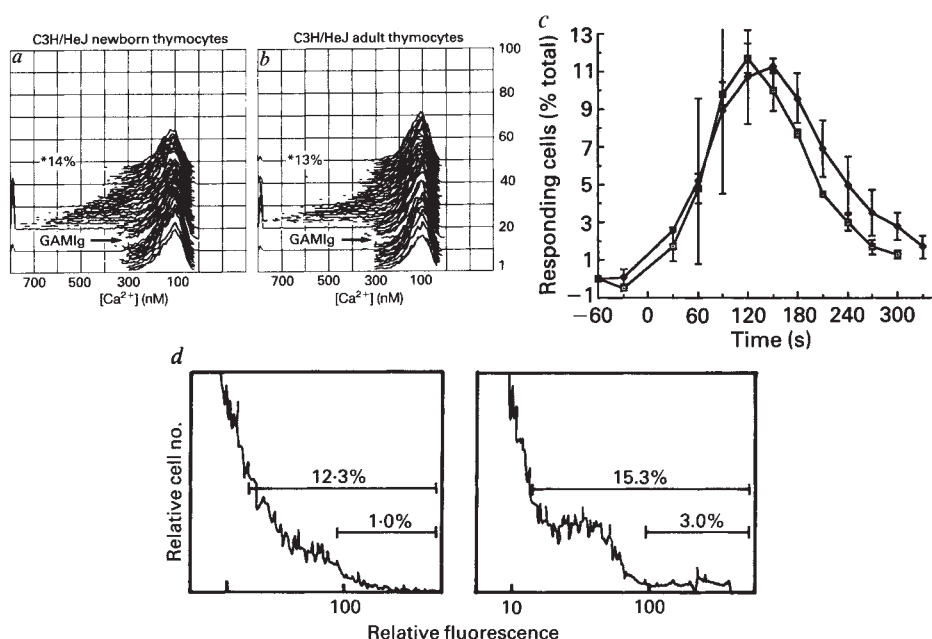
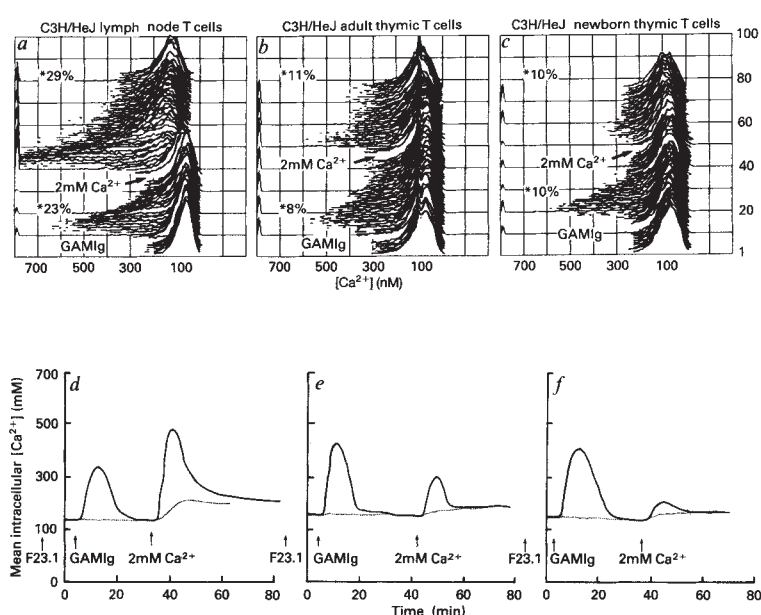


Fig. 3 Induction of Ca^{2+} mobilization in Ca^{2+} -depleted medium in response to crosslinking of T cell receptor on lymph-node T cells, adult thymic T cells and newborn thymic T cells. The graphs on the left (*a*, *b* and *c*) depict flow cytometric analysis. Asterisks indicate % of cells responding with increased $[\text{Ca}^{2+}]_i$; 7 channels represent 1 min on the timescale. The graphs on the right (*d*, *e* and *f*) depict the mean $[\text{Ca}^{2+}]_i$ of the responding populations with (—) or without (···) preincubation with F23 after addition of the crosslinking reagent. The mean $[\text{Ca}^{2+}]_i$ of the responding population was determined by plotting the mean indo-1 fluorescence ratio of cells above the threshold ratio value (as described in the legend to Fig. 1) against time. Increase in $[\text{Ca}^{2+}]_i$ in Ca^{2+} -depleted medium was assumed to be from an intracellular source. GAMlg was added at the times indicated by arrows. At the second arrows, the medium was replenished with Ca^{2+} . Increase in $[\text{Ca}^{2+}]_i$ seen after addition of Ca^{2+} to the medium were postulated to be due to influx of extracellular Ca^{2+} . *a*, C3H/HeJ LN T cells + F23 + GAMlg (resuspended in <2 nM Ca^{2+} before addition of GAMlg). *b*, C3H/HeJ adult thymic T cells + F23 + GAMlg (resuspended in <2 nM Ca^{2+} before addition of GAMlg). *c*, C3H/HeJ newborn thymic T cells + F23 + GAMlg (resuspended in <2 nM Ca^{2+} before addition of GAMlg). *d*, C3H/HeJ LN T cells + F23 (—) or -F23 (···) + GAMlg (resuspended in <2 nM Ca^{2+} before addition of GAMlg). *e*, C3H/HeJ adult thymic T cells + F23 (—) or -F23 (···) + GAMlg (resuspended in <2 nM Ca^{2+} before addition of GAMlg). *f*, C3H/HeJ newborn thymic T cells + F23 (—) or -F23 (···) + GAMlg (resuspended in <2 nM Ca^{2+} before addition of GAMlg).

Methods. Nylon wool purified lymph node cells from C3H/HeJ adult mice or unfractionated thymocytes from C3H/HeJ adult or newborn mice (day 19 of gestation) were loaded with indo-1 and incubated without or with F23, as in Fig. 1. Cells were resuspended in medium containing 0.2 mM EGTA (<2 nM Ca^{2+} by fluorimetric assay using indo-1). GAMlg was added as indicated. The medium was replenished with Ca^{2+} to a final concentration of 2 mM.



First, we repeated our experiments using cells from newborn (day 20) C3H mice. These animals contain almost normal percentages of receptor-positive immature thymocytes and virtually no mature cells¹⁷. As shown in Figs 2*d* and 2*e*, the percentage of F23^+ thymocytes was similar in the newborn and adult mice, although only 1% of cells in newborn thymus were identified as phenotypically 'mature' on the basis of a high level expression of TcR. Crosslinking of F23^+ receptors induced an increase in

$[\text{Ca}^{2+}]_i$ in both adult and newborn mice (Figs. 2*a*, 2*b* and 2*c*). Therefore, in populations $\geq 99\%$ pure for immature thymocytes, virtually all receptor-positive cells mobilize Ca^{2+} in response to crosslinking of the TcR. These data indicate that the immature thymocyte population bearing low density surface antigen receptor can be activated to mobilize Ca^{2+} by engagement of the TcR.

To further corroborate these findings, we used the monoclonal antibody KJ23a, an anti- $\alpha\beta/\text{T3}$ antibody which recognizes

V β 17a, found on T cells of the mouse strains SWR, SJL, C57/L and C57/Br (refs 18, 19). T cells expressing V β 17a are selectively eliminated from the peripheral T cell and mature thymocyte pool of mice expressing IE (such as C57/Br), but are present in expected numbers in the immature thymocyte population of such animals. Therefore, KJ23a⁺ thymocytes of strain C57/Br represent a >99.9% pure population of immature thymocytes with low receptor density. Crosslinking of KJ23a on thymocytes of strain C57/Br, of which 2-3% bind KJ23a by fluorescent staining (ref. 19, and unpublished observations for C57/Br), resulted in an increase in $[Ca^{2+}]_i$ by $1.8 \pm 0.7\%$ ($n = 12$) of cells, compared with $0.2 \pm 0.4\%$ ($n = 10$) ($P < 0.01$) of cells coated with F23. (C57/Br has deleted the allele V β 8 and therefore does not express the receptor bound by F23.) These data therefore provide further evidence that the TcR of immature thymocytes is functional in mediating calcium mobilization.

To dissect the Ca^{2+} mobilization response which follows TcR crosslinking, we used an analytical approach which allows separation of intracellular Ca^{2+} release and extracellular Ca^{2+} influx. When extracellular free Ca^{2+} was reduced from 1 mM to <2 nM with Ca^{2+} /EGTA buffered medium, crosslinking of F23 on indo-1-loaded lymph node T cells from strain C3H/HeJ resulted in a transient increase in $[Ca^{2+}]_i$ (Fig. 3a). $[Ca^{2+}]_i$ declined rapidly to the baseline $[Ca^{2+}]_i$ measured before activation. With repletion of extracellular Ca^{2+} , a second rise in $[Ca^{2+}]_i$ was seen in [Fig. 3a]. Fig. 3d shows an increase in mean $[Ca^{2+}]_i$ of the responding population, as well as the increase in basal $[Ca^{2+}]_i$ in control cells in low Ca^{2+} following repletion (dotted line). These results therefore support and extend previous reports^{20,21} and suggest that there are two components to the increase in $[Ca^{2+}]_i$ induced by crosslinking of TcR: a rise due to release from intracellular stores, and a rise which is dependent upon extracellular Ca^{2+} , presumably reflecting extracellular influx.

As shown in Figs 3b and 3c, the release of Ca^{2+} from intracellular stores in activated adult and newborn thymocytes is similar in extent and kinetics to that in activated peripheral T cells (Fig. 3a). Extracellular influx of Ca^{2+} is, however, significantly reduced in adult thymocytes compared to peripheral T cells (compare Fig. 3b with Fig. 3a) and is even further reduced in newborn thymocytes (compare Fig. 3c with Figs 3b and 3a). These data are also shown as mean $[Ca^{2+}]_i$ of the responding population (Figs 3d-f). Again, release of Ca^{2+} from intracellular stores is similar in mature and immature T cells, whereas extracellular influx of Ca^{2+} is markedly reduced in immature cells (compare Figs 3d and 3f).

The different biological consequences of receptor crosslinking in mature and immature T cells could be a result of differences in the extent of Ca^{2+} entry from the extracellular environment. Alternatively, it is possible that the Ca^{2+} second messenger is coupled to a different biological response programme in mature and immature T cells.

Received 5 May; accepted 4 September 1987.

- Marrack, P. & Kappler, J. *Immunological Disease* (ed. Samter, M.) (Little, Brown and Co., Chicago, 1985).
- McDuffie, M., Born, W., Marrack, P. & Kappler, J. *Proc. Natn. Acad. Sci. U.S.A.* **83**, 8728-8732 (1986).
- Born, W. *et al. J. Immun.* **138**, 999-1008 (1987).
- Chen, W.-F., Scollay, R., Clark-Lewin, E. & Shortman, K. *Thymus* **5**, 179 (1983).
- Chen, W.-F., Scollay, R. & Shortman, K. *J. Immun.* **129**, 18-24 (1982).
- Ceredig, R., Glasebrook, A. L. & MacDonald, H. R. *J. exp. Med.* **155**, 358-379 (1982).
- Kisielow, P., von Boehmer, H. & Haas, W. *Eur. J. Immun.* **12**, 463-467 (1982).
- Zlotnik, A., Ransom, J., Frank, G., Fischer, M. & Howard, M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3856-3860 (1987).
- Kuno, M. & Gardner, P. *Nature* **326**, 301-304 (1987).
- Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
- Meuer, S. C. *et al. J. exp. Med.* **157**, 705-719 (1983).
- Weiss, A. *et al. A. Rev. Immun.* **4**, 593-619 (1986).
- Nisbet-Brown, E., Lee, J. W. W., Cheung, R. K. & Gelfand, E. W. *Nature* **316**, 545-547 (1985).
- Ledbetter, J. A., June, C. H., Grosmaire, L. S. & Rabinovitch, P. S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1384-1388 (1987).
- Crispe, I. N., Bevan, M. J. & Staerz, U. D. *Nature* **317**, 627-629 (1985).
- Staerz, U. D., Rammensee, H.-G., Benedetto, J. & Bevan, M. J. *J. Immun.* **134**, 3994-4000 (1985).

- Roehm, N. *et al. Cell* **38**, 577-584 (1984).
- Kappler, J. W. *et al. Cell* **49**, 263-271 (1987).
- Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
- Gelfand, E. W., Mills, G. B., Cheung, R. K., Lee, J. W. W. & Grinstein, S. *Immun. Rev.* **95**, 59-87 (1987).
- Imboden, J. B. & Stobo, J. D. *J. exp. Med.* **162**, 446-456 (1985).
- Gryniewicz, G., Poenie, M. & Tsien, R. Y. *J. biol. Chem.* **260**, 3440-3450 (1985).
- Rabinovitch, P. S., June, C. H., Grossmann, A. & Ledbetter, J. A. *J. Immun.* **137**, 952-961 (1986).

Retrograde transport by the microtubule-associated protein MAP 1C

Bryce M. Paschal & Richard B. Vallee

Cell Biology Group, Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts 01545, USA

Microtubules are involved in several forms of intracellular motility, including mitosis and organelle movement. Fast axonal transport is a highly ordered form of organelle motility that operates in both the anterograde (outwards from the cell body) and retrograde (from the periphery towards the cell body) direction¹. Similar microtubule-associated movement is observed in non-neuronal cells², and might be involved in secretion, endocytosis and the positioning of organelles within the cell^{3,4}. Kinesin⁵ is a mechanochemical protein that produces force along microtubules in an anterograde direction⁶. We recently found that the brain microtubule-associated protein MAP 1C (ref. 7) is a microtubule-activated ATPase⁸ and, like kinesin, can translocate microtubules in an *in vitro* assay for microtubule-associated motility⁸. MAP 1C seemed to be related to the ciliary and flagellar ATPase, dynein, which is thought to produce force in a direction opposite to that observed for kinesin⁹. Here we report that MAP 1C, in fact, acts in a direction opposite to kinesin, and has the properties of a retrograde translocator.

We used axonemes from the biflagellate unicellular alga *Chlamydomonas*, which have served as a general standard for microtubule polarity¹⁰, to determine the polarity of force production by MAP 1C. The microtubules comprising the axoneme are all oriented with the same polarity, and the end of the axoneme that had been distal to the *Chlamydomonas* cell body *in vivo* (the + end) has a marked tendency to fray during preparation¹⁰⁻¹³. This is in contrast to the proximal end (the - end), which remains compact.

We prepared MAP 1C and kinesin (see Fig. 1) and allowed the proteins to adsorb to a coverslip. We applied salt-extracted axonemes and assayed them for gliding motility in the presence of ATP. Axonemes alone in the presence of ATP showed no movement, as we expected. When applied to MAP 1C-coated coverslips, we saw axonemal gliding (Fig. 1a). Gliding was always from the compact (proximal or -) to the frayed (distal or +) end. In contrast, kinesin caused axonemes to glide in the opposite direction (Fig. 1b).

As an additional means of determining polarity, we grew microtubules onto axonemes from purified tubulin subunits. Microtubule assembly at the distal end occurs more rapidly than at the proximal end, resulting in longer microtubules at the distal end¹⁰⁻¹³. In preliminary experiments, we saw some variation in the relative lengths of microtubules at the two ends. This could be controlled by maintaining a constant concentration of free tubulin during the manipulations involved in applying the axonemes to the coverslip. Examples of axonemes with microtubules assembled at the ends are shown in Fig. 2. Long microtubules can be seen at the frayed end of the axoneme, and short or no microtubules at the proximal end. In the example shown for MAP 1C (Fig. 2a), the entire complex has moved from the short-microtubule to the long-microtubule end. This corresponds to the direction of movement seen with the

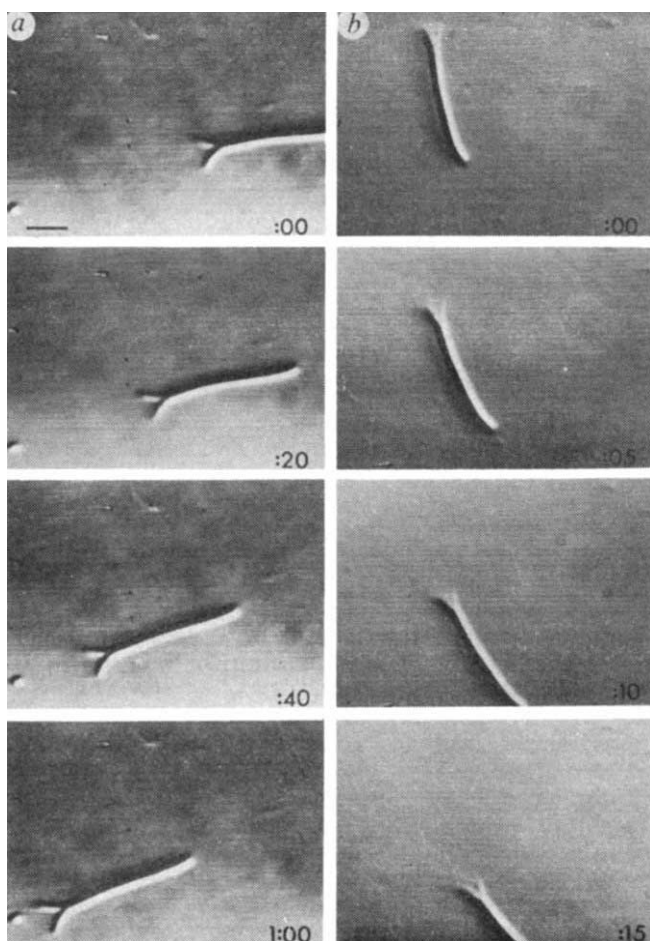


Fig. 1 Gliding polarity determined using *Chlamydomonas* axonemes. MAP 1C (*a*) or kinesin (*b*) was adsorbed to a coverslip, and *Chlamydomonas* flagellar axonemes applied. The axonemes glided toward their frayed (distal or +) end on MAP 1C, and towards their compact (proximal or -) end on kinesin. Scale bar, 2 μm .

Methods. MAP 1C was obtained from calf brain white matter and purified by sucrose density-gradient centrifugation⁸. Kinesin was purified from the same source using a modification of the method of Vale *et al.*⁵. Following AMPPNP-induced binding to microtubules, kinesin was extracted with 10 mM ATP. The extract was applied to a 5–20% linear sucrose gradient. A peak of kinesin at 10S was detected using gel electrophoresis (see ref. 8), and this material was used undiluted for the motility assay. This method provided better separation of kinesin from contaminating MAP 1C than the chromatographic procedures that we tried. Axonemes were prepared from *Chlamydomonas reinhardtii* and extracted with high salt to remove the outer dynein arms, as described previously²³. Motility was assayed as described previously⁸ using axonemes instead of taxol-stabilized brain microtubules. In brief, 4 μl of MAP 1C or kinesin was applied to a coverslip and the protein was allowed to adsorb for 5–10 min. Flagellar axonemes (2 μl) were added together with MgATP (final concentration 2 mM) and Tris/KCl buffer (20 mM Tris/HCl, pH 7.6, containing 50 mM KCl, 5 mM MgSO_4 and 0.5 mM EDTA) to a final volume of 10 μl . Motility was examined at ambient temperature (25–30 °C) by video-enhanced differential interference-contrast microscopy⁸. Time is indicated in each panel in min. Average gliding rate for axonemes on MAP 1C was 0.2 $\mu\text{m s}^{-1}$, slower than for free microtubules⁸. A similar difference has also been observed for kinesin¹⁵.

axonemes alone (Fig. 1*a*). We saw many examples in which the structure was torn apart by the microtubules grown onto the frayed end. An early stage of this behaviour is shown in Fig. 2*a* (see final panel). In other cases, the axoneme remained stationary, but the microtubules at the proximal end became

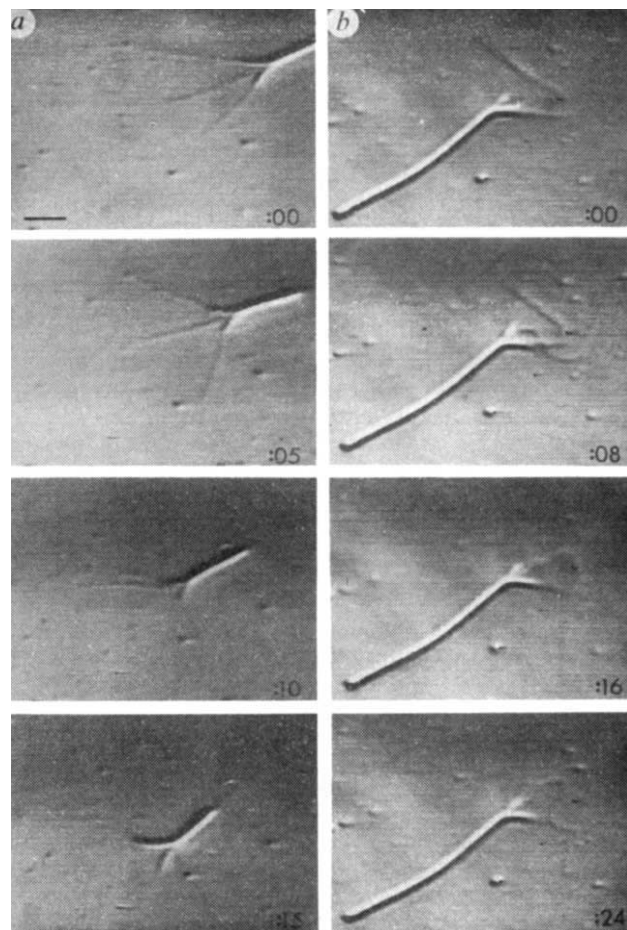


Fig. 2 Polarity determined using microtubules assembled onto flagellar axonemes. *a*, MAP 1C; *b*, kinesin. Purified calf brain tubulin⁸ was allowed to assemble onto *Chlamydomonas* axonemes. As reported previously^{10–13}, the microtubules at the frayed (distal or +) end of the axoneme were longer than those at the compact (proximal or -) end. *a*, Gliding motility of the entire microtubule/axoneme complex. Short microtubules seen at the proximal end. *b*, Axoneme remaining stationary: microtubules seen only at the distal end and showed writhing movement, characteristic of microtubules experiencing a force but tethered at one end^{6,14}.

Methods. The motility assay and preparative methods were as described in Fig. 1. Purified tubulin⁸ was added to a final concentration of about 1.5 mg ml^{-1} to flagellar axonemes in Tris/KCl buffer + 1 mM GTP and incubated for 5 min at 37 °C. Tubulin was added to the same final concentration to the sample of MAP 1C or kinesin already applied to the coverslip, and the axoneme/microtubule complex was then added and monitored for motility.

contorted, whereas those at the distal end remained straight (compare with Fig. 2*b*, and see below).

In experiments performed with kinesin, using axonemes as seeds for brain microtubule growth, we again found the direction of movement and force production to be opposite to that observed with MAP 1C. Gliding of axonemes with associated microtubules was always toward the shorter microtubules. In cases in which the axoneme remained stationary (Fig. 2*b*), the behaviour of the microtubules at the ends was opposite to that observed for MAP 1C (see above). With kinesin, the microtubules assembled at the distal end showed contorted bending and whiplashing movements. This behaviour occurs when a microtubule is tethered at one end^{6,14}, in this case the end at which the assembled microtubule is attached to the axoneme. Force produced along the microtubule will cause it to bend as it is incapable of net translocation across the coverslip. In

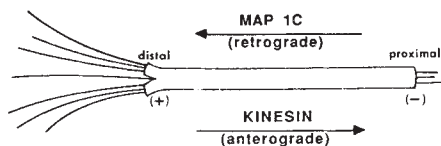


Fig. 3 Diagrammatic representation of force production. A *Chlamydomonas* axoneme is shown with long microtubules assembled onto frayed (distal or +) end, and short microtubules assembled onto compact (proximal or -) end¹⁰⁻¹³. Axonemes were driven towards their distal ends by MAP 1C (Figs 1, 2). Thus, a cytoplasmic structure attached to MAP 1C would be driven in the opposite direction, or towards the proximal end of the microtubule. In axons this would correspond to retrograde motility, as axonal microtubules have the same polarity as those in the *Chlamydomonas* axoneme relative to the cell body^{24,25}. Axonemes and microtubules assembled at their ends were driven toward the proximal end of the axoneme by kinesin. This would correspond to anterograde motility for an axonal organelle.

experiments using kinesin (Fig. 2b), the force appeared to be directed towards the proximal end of the axoneme, the same direction observed for gliding of the axonemes alone (Fig. 1b). We observed no microtubules at the proximal end in the example shown in Fig. 2b, but when they were present, they remained extended, the behaviour expected for microtubules pulled away from their point of attachment^{6,14}. As noted above, we saw the opposite behaviour with MAP 1C.

Our results indicate that MAP 1C produces force in the direction expected for a retrograde translocator (Fig. 3), in the opposite direction to that mediated by kinesin, which we confirm here acts as an anterograde translocator^{6,15}. To obtain additional insight into the possible role of MAP 1C in organelle transport

in vivo, and for further comparison with other motile factors, we examined the effect of some pharmacological agents and nucleotides on microtubule-gliding motility (Table 1). Only ATP supported MAP 1C-mediated motility, in contrast with kinesin, which can use GTP and ITP, as well as ATP¹⁵. MAP 1C motility was also completely inhibited by low levels of *N*-ethylmaleimide (NEM), to which kinesin is notably insensitive^{5,15}. The non-hydrolysable ATP analogue AMPPNP (adenyl-5'-yl imidodiphosphate) inhibited MAP 1C motility at levels similar to those reported to for kinesin^{5,15}. These levels have also been reported to inhibit general vesicle motility in extruded squid axoplasm¹⁶, although in that study no distinction was made between the sensitivity of anterograde and retrograde transport. Vanadate inhibited MAP 1C motility at concentrations as low as 10 μ M, whereas kinesin is less sensitive to this reagent¹⁵. Again, general organelle transport has been reported to be inhibited by vanadate in this concentration range¹⁷, but the sensitivity of anterograde and retrograde motility have not been compared. Note that the sensitivity of MAP 1C both to vanadate and to NEM corresponded to the characteristics of the residual retrograde activity observed by Vale *et al.* in crude squid cytosol following removal of kinesin⁶.

One agent that has been reported to inhibit retrograde organelle transport specifically *in vivo* is the adenosine analogue EHNA (erythro-9-[3-(2-hydroxypropyl)] adenine; ref. 18). MAP 1C motility was completely inhibited by EHNA, suggesting MAP 1C is involved in retrograde organelle transport. EHNA has also been reported to be a specific inhibitor of the axonemal ATPase, dynein¹⁹. In fact, the limited nucleotide specificity of MAP 1C, and its sensitivity to vanadate, AMPPNP and NEM are all characteristic of axonemal dynein¹⁹⁻²¹. As noted above, dynein is also thought to exert a force along microtubules within the axoneme that is directed from the proximal to the distal end of the structure⁹. Thus, our present findings are consistent with the hypothesis that MAP 1C is a soluble form of dynein⁸ and is responsible for retrograde organelle transport in cells. We can now isolate MAP 1C from rat liver as well as from brain (C. Collins and R.B.V., unpublished), indicating a role in diverse cell types.

By analogy with axonemal dynein, MAP 1C could also be involved in force production between adjacent cytoplasmic microtubules. Such behaviour could play a part in the reorganization of the cytoskeleton²² or in mitosis. However, based on the high concentration of MAP 1C in brain, we currently favour a role in retrograde organelle transport.

We thank Drs Christine Collins, Howard Shpetner and Nancy Kravit for critical reading of this manuscript, and Drs Stephen King and George Witman for supplying *Chlamydomonas* axonemes.

Received 21 August; accepted 29 September 1987.

Table 1 Pharmacological properties and nucleotide specificity of MAP 1C motility

Nucleotide	Concentration	Motility
MgATP	0.010–5 mM	+
MgGTP	5 mM	–
MgCTP	5 mM	–
MgTTP	5 mM	–
MgUTP	5 mM	–
MgITP	5 mM	–
MgPPP _i	5 mM	–
Inhibitor		
NEM*	0.1 mM	–
AMPPNP†	1 mM	±
AMPPNP†	2 mM	–
Sodium vanadate	10 μ M	±
Sodium vanadate	50 μ M	–
EHNA‡	1 mM	–
Sodium azide	5 mM	+
Ouabain	0.05 mg ml ⁻¹	+

Motility assays as described for Fig. 1. All stock nucleotide solutions contained equimolar MgSO₄. Inhibitor effects were examined in the presence of 2 mM MgATP unless otherwise noted. +, Good motility, $\geq$ 90% of microtubules showing unidirectional gliding; $\pm$, a few gliding or whiplashing microtubules, or variable results on repeated trials; –, no detectable motility.

* MAP 1C was incubated with NEM for 15 min at room temperature. The reaction was quenched with 10 mM dithiothreitol (DTT) before adsorption of the protein onto the coverslip (see ref. 15). DTT alone had no effect on motility.

† Performed in the presence of 1 mM ATP.

‡ EHNA in tenfold excess over ATP (0.1 mM). No inhibition was observed with equimolar EHNA and ATP, as expected for a dynein-like ATPase¹⁹.

- Grafstein, B. & Forman, D. S. *Physiol. Rev.* **60**, 1167–1283 (1980).
- Schliwa, M. in *Cell and Muscle Motility* Vol. 5 (ed. Shay, J.) 1–82 (Plenum, New York, 1984).
- Herman, B. & Albertini, D. F. *J. Cell Biol.* **98**, 565–576 (1984).
- Rogalski, A. A. & Singer, S. J. *J. Cell Biol.* **99**, 1092–1100 (1984).
- Vale, R. D., Reese, T. S. & Sheetz, M. P. *Cell* **42**, 39–50 (1985).
- Vale, R. D. *et al. Cell* **43**, 623–632 (1985).
- Bloom, G. S., Schoenfeld, T. A. & Vallee, R. B. *J. Cell Biol.* **98**, 320–330 (1984).
- Paschal, B. M., Shpetner, H. S. & Vallee, R. B. *J. Cell Biol.* **105**, 1273–1282 (1987).
- Sale, W. S. & Satir, P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2045–2059 (1977).
- Bergen, L. G. & Borisy, G. G. *J. Cell Biol.* **84**, 141–150 (1980).
- Allen, C. & Borisy, G. G. *J. molec. Biol.* **90**, 381–402 (1974).
- Binder, L. I., Dendler, W. L. & Rosenbaum, J. L. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1122–1126 (1975).
- Witman, G. B., Cleveland, D. W., Weingarten, M. D. & Kirschner, M. W. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4070–4074 (1976).
- Allen, R. D. *et al. J. Cell Biol.* **100**, 1736–1752 (1985).
- Porter, M. E. *et al. J. biol. Chem.* **262**, 2794–2802 (1987).
- Lasek, R. J. & Brady, S. T. *Nature* **316**, 645–647 (1985).
- Forman, D. S., Brown, K. J. & Livengood, D. R. *J. Neurosci.* **3**, 1279–1288 (1983).
- Forman, D. S., Brown, K. J. & Promersberger, M. E. *Brain Res.* **272**, 194–197 (1983).
- Penningroth, S. M. *Meth. Enzym.* **134**, 477–487 (1986).
- Warner, F. D. & Mitchell, D. R. *Int. Rev. Cytol.* **66**, 1–42 (1980).
- Johnson, K. A. A. *Rev. Biophys. biophys. Chem.* **14**, 161–188 (1985).
- Allen, R. D., Allen, N. S. & Travis, J. L. *Cell Motil.* **1**, 291–302 (1981).
- King, S. M., Otter, T. & Witman, G. B. *Meth. Enzym.* **134**, 291–306 (1986).
- Heidemann, S. R., Landers, J. M. & Hamburg, M. A. *J. Cell Biol.* **91**, 661–665 (1981).
- Burton, P. R. & Paige, J. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3269–3273 (1981).

Relation of HTLV-4 to simian and human immunodeficiency-associated viruses

Beatrice H. Hahn*†, Lilly I. Kong*, Shei-Wen Lee*,
Prasanna Kumar*, Maria E. Taylor*,
Suresh K. Arya§ & George M. Shaw*‡

Departments of * Internal Medicine, † Microbiology and

‡ Biochemistry, University of Alabama at Birmingham,
Birmingham, Alabama 35294, USA

§ Laboratory of Tumor Cell Biology, National Institutes of Health,
Building 37, Bethesda, Maryland 20895, USA

Human immunodeficiency virus type 1 (HIV-1) is the aetiological agent of AIDS (acquired immune deficiency syndrome) in most countries¹⁻³ and probably originated in Central Africa like the AIDS epidemic itself. Evidence for a second major group of human immunodeficiency-associated retroviruses came from a report that West African human populations⁴ like wild-caught African green monkeys⁵ had serum antibodies that reacted more strongly with a simian immunodeficiency virus (STLV-3_{Mac}) (ref. 6) than with HIV-1. Novel T-lymphotropic retroviruses were reported to have been isolated from healthy Senegalese West Africans (HTLV-4) (ref. 4) and from African green monkeys (STLV-3_{AGM}) (ref. 7), and a different retrovirus (HIV-2) was identified in other West African AIDS patients⁸. Genomic analysis of HIV-2 clearly distinguished it from STLV-3 (ref. 9), but restriction enzyme site-mapping of three different HTLV-4 isolates and six different STLV-3_{AGM} isolates showed them to be essentially indistinguishable¹⁰⁻¹². In this report we clone, restriction map, and partially sequence three isolates of HTLV-4 (PK82, PK289, PK190) (ref. 4). We find that these viruses differ in nucleotide sequence from each other and from three isolates of STLV-3_{AGM} (K78, K6W, K1) (ref. 7) by 1% or less. We also report the isolation of a T-lymphotropic retrovirus from the peripheral blood of a healthy Senegalese

woman which hybridizes preferentially to HIV-2 specific DNA probes. We conclude that HTLV-4 (ref. 4) and STLV-3_{AGM} (ref. 7) are not independent virus isolates and that HIV-2 is present in Senegal as it is in other West African countries¹³.

Because of similarities in structure and antigenicity among HIV-1, HIV-2 and STLV-3 (also referred to as SIV, simian immunodeficiency virus), some relation might be expected between their nucleic acid sequences and phylogenesis. Restriction enzyme analysis showed that six out of six STLV-3_{AGM} isolates and two out of three HTLV-4 isolates were identical in 32 out of 32 restriction sites mapped¹⁰⁻¹². The third HTLV-4 isolate was identical to the others in 31 out of 32 mapped restriction sites. Of three STLV-3_{Mac} isolates obtained independently at the New England Regional Primate Center, one (isolate 251) was identical to the STLV-3_{AGM}/HTLV-4 consensus pattern in 23 out of 23 sites, and the others were identical to each other and similar to the consensus pattern in 21 out of 23 sites¹¹. The same genetic analysis clearly distinguished HIV-2 from these viruses⁹.

To define more precisely the relation between different HTLV-4, STLV-3, HIV-2 and HIV-1 viruses, a combination of restriction endonuclease cleavage mapping, molecular cloning, and nucleotide sequencing was used to analyse and compare different isolates of HTLV-4 and STLV-3. The DNA from two HTLV-4 isolates (HTLV-4/PK82, HTLV-4/PK289) (ref. 4) and two STLV-3 isolates (STLV-3_{AGM}/K1, STLV-3_{AGM}/Mm 6324) (ref. 7) was extracted from virally-infected cell pellets without interim cell cultivation. Two other virus isolates (HTLV-4/PK190, STLV-3_{AGM}/K1) (ref. 4, 7) were cultured and used to prepare RNA for a viral-specific partial complementary DNA clone. A viral-specific insert from one of these clones (HTLV-4/PK190, clone pV2) was used to probe recombinant lambda phage libraries prepared from DNA infected with HTLV-4/PK82 and HTLV-4/PK289. Full-length proviral DNA clones, including flanking cellular sequences, were obtained for both HTLV-4 viruses. The restriction enzyme cleavage patterns of STLV-3_{AGM}/K1, STLV-3_{AGM}/Mm6324, HTLV-4/PK82, HTLV-4/PK289 and of the full-length lambda phage clones derived

Table 1 Nucleotide sequence differences between molecular clones of HTLV-4, STLV-3_{AGM} and STLV-3(SIV)_{Mac}

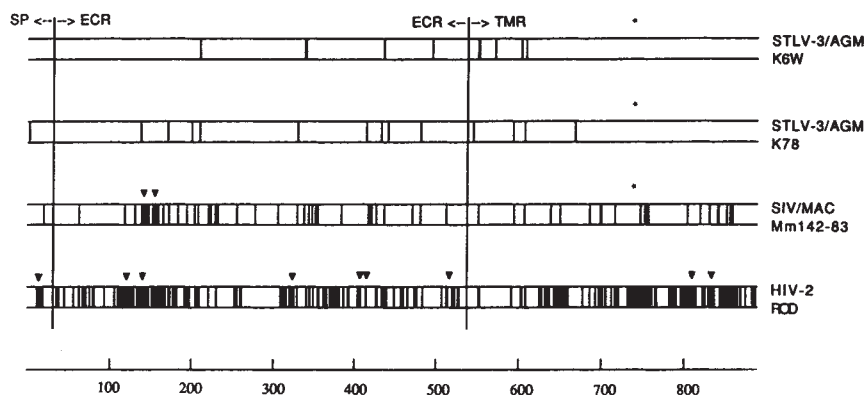
HTLV-4/PK82 provirus λ PKE102	LTR	gag	pol	sor	X	R	tat	art	env	3' orf	total
compared with:											
HTLV-4/PK190 cDNA pV2	4/519 0.8%									3/736 0.4%	6/850 0.7%
STLV-3 _{AGM} /K78 provirus λ K2-10	13/807 1.6%			5/227 2.2%	7/336 2.1%	2/291 0.7%	1/393 0.3%	0/324 0%	21/2643 0.8%	6/789 0.8%	48/4520 1.1%
STLV-3 _{AGM} /K6W provirus	13/796 1.6%	7/637 1.1%		1/227 0.4%	—† —	—† —	1/352† 0.3%	0/324 0%	16/2643 0.6%	6/776 0.7%	44/5036 0.9%
STLV-3 _{AGM} /K1 cDNA pBS3	1/668 0.1%								2/280 0.7%	3/780 0.4%	4/1160 0.4%
STLV-3 _{AGM} /K6W provirus											
compared with:	30/780 3.9%	43/1518 2.8%	95/3168* 3.0%	26/642* 4.0%	(8/336)‡ (2.4%)	(16/291)‡ (5.5%)	27/352† 7.6%	19/324 5.9%	124/2643 4.7%	62/776 7.7%	351/9090 3.9%
SIV _{Mac} /Mm142-83 provirus SIV-1											

Alignments were performed using the program NUCALN (ref. 21). The number of base-pair changes per number of base pairs compared and the percentage base-pair differences in the aligned open reading frames are shown. Note that values are given for each open reading frame separately and, because some reading frames overlap, that the total number of changes indicated in the last column is slightly less than the sum of the differences for all reading frames. The nucleotide sequences of STLV-3_{AGM}/K6W, STLV-3_{AGM}/K78 and SIV_{Mac}/Mm142 have been published^{12,14,15}.

* Nucleotide sequence alignment in the *pol* and *sor* open reading frames of K6W and Mm142-83 shows a total of nine single base-pair deletions followed 4 to 61 base pairs downstream by an equal number of single base-pair insertions. These probably represent sequencing errors because the resulting frame shifts would cause major amino-acid sequence changes in otherwise highly related proteins.

† STLV-3_{AGM}/K6W contains a 343bp deletion which includes the X, R and part highly of *tat* open reading frames.

‡ As X and R sequences are absent in STLV-3_{AGM}/K6W (ref. 15), nucleotide sequence information for HTLV-4/PK82 clone PKE102 is shown.



from the two HTLV-4 isolates were compared. Eight restriction enzymes (*NdeI*, *PstI*, *XbaI*, *BamHI*, *HindIII*, *SstI*, *BglII*, *EcoRI*) were used. The restriction cleavage patterns of these isolates and proviral clones (data not shown) were identical in 19 out of 19 sites mapped, and corresponded to the HTLV-4/STLV-3_{AGM} consensus map¹¹.

The DNA from proviral or cDNA clones of HTLV-4/PK82, HTLV-4/PK190 and STLV-3_{AGM}/K1 was sequenced to define their relation to each other and to HIV-2 (ref. 9) and other STLV-3 isolates^{12,14,15}. For HTLV-4/PK82, the regions sequenced included the entire long terminal repeat (LTR), *gag* p16 equivalent, central region, envelope, and 3' open reading frame (5,391 base pairs). For HTLV-4/PK190 and STLV-3_{AGM}/K1, the sequence included R and U3 LTR regions and portions of 3' *orf* (850 and 1,160 base pairs, respectively). A summary of this nucleotide sequence information is compared with published sequences of other STLV-3 viruses in Table 1. (The complete sequence information has been submitted to EMBL/GenBank under accession code number Y00283.)

From this comparison of the sequences of HTLV-4/PK190 and three different STLV-3_{AGM} (K78, K6W, K1), none differ from HTLV-4/PK82 by more than 1.1%. In the envelope gene, which varies most widely in independent isolates of HIV-1, the three STLV-3_{AGM} sequences vary from HTLV-4/PK82 by 0.8% or less. But the virus STLV-3_{Mac}/Mm142-83 differs from both HTLV-4/PK82 and STLV-3_{AGM}/K6W by about 4.7% in envelope. Figure 1 depicts the distribution of deduced amino acid changes present in the envelopes of different STLV-3_{AGM}, STLV-3_{Mac} and HIV-2 viruses with respect to HTLV-4/PK82. Compared to HTLV-4/PK82, there were 10 amino acid differences in STLV-3_{AGM}/K6W, 14 in STLV-3_{AGM}/K78, and 65 in STLV-3_{Mac}/Mm142-83. All three STLV-3 and both HTLV-4 sequences (Fig. 1 and ref. 12) contain an identical stop codon in the envelope transmembrane coding region. Also, STLV-3_{Mac}/Mm142-83 compared to HTLV-4/PK82 contains clustered nucleotide and amino acid substitutions and a three-base-pair (bp) insertion. In the same regions of the viral envelope, HTLV-4 and STLV-3_{AGM} are identical. As a result of the nearly identical restriction patterns and nucleotide sequences of three different HTLV-4 and six different STLV-3_{AGM} isolates, and because of the failure to re-isolate such STLV-3-like viruses from humans, and in view of our own isolation of an HIV-2-related virus from a healthy Senegalese woman (see below), we conclude that HTLV-4 (ref. 4) and STLV-3_{AGM} (ref. 7) are not independent laboratory isolates, but instead represent transmission of a single laboratory virus strain to multiple cultures. Furthermore, probably the various HTLV-4 and STLV-3_{AGM} isolates originated from cultures of STLV-3_{Mac} isolate 251, because the HTLV-4/STLV-3_{AGM} restriction site consensus pattern coincides precisely with isolate 251 in 32 out of 32 mapped sites, but is similar to STLV-3_{Mac} isolate 142 in only 27 out of 31 sites^{11,14,16}.

On the basis of clinical¹³ and seroepidemiologic¹⁷ studies of

Fig. 1 Distribution of amino-acid sequence differences in the envelope genes of STLV-3_{AGM}, STLV-3(SIV)_{Mac} and HIV-2 compared to HTLV-4. Vertical lines represent the position of single amino-acid sequence differences in the envelopes of STLV-3_{AGM}/K6W (ref. 15), STLV-3_{AGM}/K78 (ref. 12), SIV_{Mac}/Mm142-83 (ref. 14), and HIV-2_{ROD} (ref. 9) compared with HTLV-4/PK82, as determined by pairwise alignment using the program PRTALN (ref. 21). Arrows depict the position of amino-acid insertions and deletions. An asterisk depicts the location of a TAG in-frame stop codon in the transmembrane region which is present in the clones of HTLV-4, STLV-3_{AGM} and STLV-3(SIV)_{Mac}, but not in HIV-2_{ROD}. SP, signal peptide; ECR, extracellular region; TMR, transmembrane region. Relative amino-acid positions are shown on the bottom scale.

West African populations, it is clear that a virus, or family of viruses, which is antigenically more similar to HIV-2 and STLV-3(SIV) than to HIV-1, is prevalent and in some cases associated with immunodeficiency. There is no confirmed evidence that STLV-3(SIV) is present in these human populations, whereas HIV-2 has been isolated from patients in Cape Verde and Guinea Bissau¹³. We attempted to isolate retroviruses from the peripheral blood of four healthy seropositive Senegalese prostitutes. Mononuclear cell preparations were co-cultivated with uninfected normal donor lymphocytes. After 1-2 weeks,

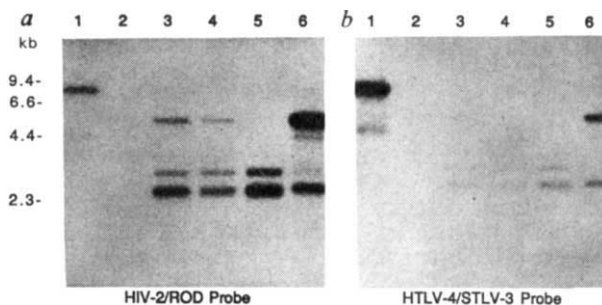


Fig. 2 Determination of relative nucleotide sequence homology of a newly isolated Senegalese retrovirus (isolate ST; see text) to HIV-2 (panel a) and HTLV-4/STLV-3 (panel b). Panels a and b are identical except for the viral-specific nucleic acid probe used. Lane 1, HTLV-4/PK82 (ref. 4); Lane 2, normal human DNA; Lanes 3-6, 4 different subcultures derived from the parental ST virus isolate (see Methods). Preferential hybridization of the HIV-2 specific probe to ST isolate DNA is seen (panel a, lanes 3-6). The HTLV-4/STLV-3 specific probe hybridizes much less efficiently to ST DNA (panel b, lanes 3-6) but hybridizes to HTLV-4 DNA on the same filter (panel b, lane 1). Lanes 5 and 6 (panel a) show two distinct ST viral genotypes polymorphic in a single *Bam*HI restriction site. Virus cultures (lanes 3 and 4) contain a mixture of these two predominant viral forms which also differ in their *Xba*I restriction patterns (data not shown). Size markers are given on the left in kilobases (kb).

Methods. Virus was isolated by co-cultivation of patient ST and normal donor peripheral blood mononuclear cells, followed by transmission of virus to Hut78 cells (lanes 3-5) or SupT1 cells (lane 6). Lane 5 shows a virus culture derived by limiting dilution cell cloning of the isolate in lane 4. Total cellular and viral DNA was extracted from infected and uninfected cells as described¹⁸. 10 µg of each DNA was digested with *Bam*HI, separated by 0.8% agarose gel electrophoresis and transferred to 2 nitrocellulose filters; one filter (panel a) was hybridized to a 4.3-kb-HIV-2-specific probe from pROD35 (ref. 9) and the other (panel b) to a corresponding 7.5-kb-HTLV-4/STLV-3-specific probe from λPKE102 (see text and Table 1). Hybridization was at 37 °C for 18 h with probe at 10×10^6 c.p.m. ml⁻¹ (specific activity of probe was $2-3 \times 10^8$ c.p.m. µg⁻¹ DNA) in 2.4×SSC, 40% formamide and 10% dextran sulphate, as described¹⁸. Filters were washed in 1×SSC + 0.2% sodium lauryl sulphate at 65 °C for 3 h and exposed to Kodak XAR-5 film for 4 h.

transient reverse transcriptase activity was detected in culture supernatants from three individuals, and virus from one of these cultures (isolate ST) was transmitted successfully to four immortalized T-cell lines, SUPT1, HUT78, H9, and CEM174. On the basis of its ultrastructural morphology, particulate reverse transcriptase activity, T4 cell tropism and antigenic reactivity by Western blot hybridization analysis, this virus was related to both HIV-2 and STLV-3(SIV) (manuscript in preparation). Southern blots of virally-infected cellular DNA from these ST cultures using DNA probes specific for HIV-2 and for HTLV-4/STLV-3, however, clearly showed the ST isolate to be more related to HIV-2 than to STLV-3/HTLV-4, as shown in Fig. 2. The proviral DNA of the Senegalese isolate hybridized preferentially to HIV-2 and the restriction pattern of ST was unique. Moreover, two related yet distinguishable viral genotypes were found in the ST isolate, reflecting a genotypic heterogeneity characteristic of both HIV-1 and HIV-2 (refs 9, 18–20). The STLV-3(SIV) virus was not detected in any of the four Senegal cultures. These data demonstrate the existence of HIV-2 in Senegal and suggest that West African individuals seropositive for HTLV-4 (refs 4, 17) are indeed infected with a retrovirus, but that this belongs to the HIV-2, not STLV-3(SIV) group. This does not preclude the existence in human populations in West Africa or elsewhere of retroviruses distinct from HIV-1 and HIV-2, which is quite probable given their highly mutable nature and biological properties.

We thank P. Kanki and M. Essex for providing virus cultures of HTLV-4 and STLV-3, as well as samples of fresh peripheral blood from subjects in Senegal. This work was supported in part by grants from the National Institutes of Health, the United States Army Medical Research Acquisition Activity, and the Life and Health Insurance Medical Research Fund. B.H.H. is a Special Fellow of the Leukemia Society of America and G.M.S. is a Pew Scholar in the Biomedical Sciences.

Received 1 July; accepted 21 September 1987.

1. Barre-Sinoussi, F. *et al. Science* **220**, 868–871 (1983).
2. Popovic, M. *et al. Science* **224**, 497–499 (1984).
3. Gallo, R. C. *et al. Science* **224**, 500–503 (1984).
4. Kanki, P. J. *et al. Science* **232**, 238–242 (1986).
5. Kanki, P. J. *et al. Lancet* **i**, 1330–1332 (1985).
6. Daniel, M. D. *et al. Science* **228**, 1201–1204 (1985).
7. Kanki, P. J., Alroy, J. & Essex, M. *Science* **230**, 951–954 (1985).
8. Clavel, F. *et al. Science* **233**, 343–346 (1986).
9. Guyader, M. *et al. Nature* **326**, 662–669 (1987).
10. Hirsch, V. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9754–9758 (1986).
11. Kornfeld, H. *et al. Nature* **326**, 610–613 (1987).
12. Hirsch, V., Riedel, N. & Mullins, J. I. *Cell* **49**, 307–319 (1987).
13. Clavel, F. *et al. New Engl. J. Med.* **316**, 1180–1185 (1987).
14. Chakrabarti, L. *et al. Nature* **328**, 543–547 (1987).
15. Franchini, G. *et al. Nature* **328**, 539–543 (1987).
16. Desrosiers, R. C. *et al. Nature* **327**, 107 (1987).
17. Kanki, P. J. *et al. Science* **236**, 827–831 (1987).
18. Shaw, G. M. *et al. Science* **226**, 1165–1171 (1984).
19. Wong-Staal, F. *et al. Science* **229**, 759–762 (1985).
20. Hahn, B. H. *et al. Science* **232**, 1548–1553 (1986).
21. Wilbur, W. J. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 726–730 (1983).

Mutations of the Kirsten-*ras* proto-oncogene in human preleukaemia

Edison Liu*†§, Brian Hjelle*, Rodman Morgan‡, Frederick Hecht‡ & J. Michael Bishop*

* Department of Microbiology and Immunology, and G. W. Hooper Foundation; † Cancer Research Institute, University of California at San Francisco, San Francisco, California, USA

‡ The Genetics Center and Cancer Center of Southwest Biomedical Research Institute, Scottsdale, Arizona, USA

The myelodysplastic syndrome (MDS) or preleukaemia is a haematological disorder characterized by low blood counts, bone marrow cells of abnormal appearance and progression to acute leukaemia in as many as 30% of patients¹. The distinctive preleukaemic and leukaemic phases of this disease make it an attractive model for neoplastic progression in human tumours. We reasoned that, because dominantly transforming genes (such as mutant alleles of *ras* proto-oncogenes) are found so frequently in acute leukaemia², the search for these genetic lesions during the clinical course of patients with MDS might give us insight into the function of oncogenes in leukaemogenesis. We report here that bone marrow cells from two of four patients with preleukaemia, and from one patient who progressed to acute leukaemia from MDS, contained a transforming allele of the Ki-*ras* proto-oncogene. In one preleukaemic patient, a novel mutation in codon 13 of this *ras* gene was detected in bone marrow cells harvested 1.5 years before the acute leukaemia developed. Our findings provide evidence that *ras* mutations may be involved in the early stages of human leukaemia.

Our most extensively studied patient, Ve, is a 30-year-old male with MDS whose initial bone marrow cytogenetics showed an inversion of chromosome 3, inv(3)(q21q26.2), a lesion associated with preleukaemic states³. One and a half years after the diagnosis of MDS, the patient developed acute leukaemia. DNA

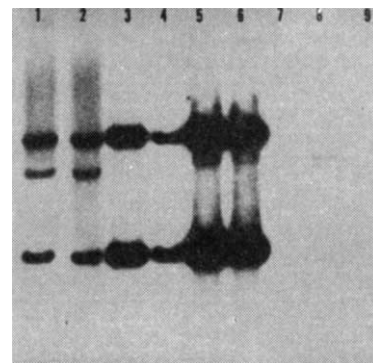


Fig. 1 Transfer of the Ki-*ras* gene with the transforming potential of DNA from a patient with myelodysplasia. DNAs from patient material and nude-mouse tumours were analysed by Southern blot hybridization for the presence of the human Ki-*ras* gene. The probe used was the 380-bp *Eco*RI-*Xba*I insert from pSVEKis encompassing exons 1, 2 and part of 3 of the human Ki-*ras*2 gene⁸. Sources of DNA: lane 1, Ve lymphoblastoid cell line; lane 2, Ve preleukaemic bone marrow; lane 3, a primary nude-mouse tumour T11A that arose after transfection with Ve preleukaemic bone marrow DNA; lane 4, a second primary nude-mouse tumour, T12A, from transfection with Ve preleukaemic DNA; lane 5, a secondary nude-mouse tumour, T10L, which arose from a transfection with T11A DNA; lane 6, secondary nude-mouse tumour, T10R, which arose from a second transfection with T11A DNA; lanes 7–9, nude-mouse tumours bearing no Ki-*ras*.

was isolated from the patient's bone marrow at the time of diagnosis of MDS and saved for subsequent analysis. A cytogenetically normal lymphoblastoid cell line was also established by infection of peripheral blood leukocytes from Ve with Epstein-Barr virus.

We used a sensitive tumorigenicity assay using athymic nude mice to identify transforming genes in DNA (refs 4 and 5; table 1). Control transfections using normal human DNA gave one tumour in 13 transfections, with a latency of 50 days. By contrast, transfection with DNA from the T24 bladder carcinoma cell line (which carries a mutant allele of the cellular Harvey-*ras* oncogene, c-Ha-*ras*) gave tumours in seven of seven transfections, with a latency of 14–21 days. Transfections using

§ Present address: Lineberger Cancer Research Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 27514, USA.

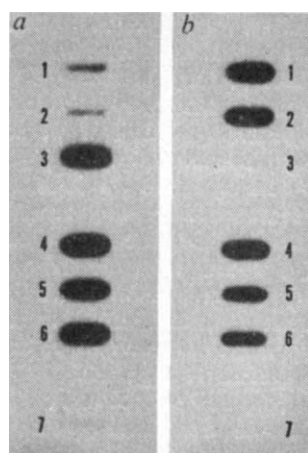


Fig. 2 Identification of the codon 13 *Ki-ras* mutation in sequential bone marrow and blood samples from a patient with myelodysplasia. *a*, Filter hybridized with oligomer probe specific for the codon-13 *Ki-ras* mutation: TTGCCTACGTCACCAAGCTCC; *b*, filter hybridized with oligomer probe specific for the wild-type *Ki-ras* allele: TTGCCTACGCCACCAAGCTCC. DNAs tested were from: 1, K562 cells; 2, Ve lymphoblastoid cell line; 3, nude-mouse tumour T12A; 4, Ve preleukaemic bone marrow, July 29 1985; 5, Ve buffy coat, May 29 1986; 6, Ve buffy coat, January 26 1987 (conversion to acute leukaemia); 7, nude-mouse tumour T12A, no amplification (no polymerase added).

Methods. The polymerase chain reaction¹¹ was used to amplify the region of codon 13 in *Ki-ras* using the *Taq* polymerase (R. Saiki, Cetus). The primers used in the reaction were: 5' primer, ATGACTGAATATAAACTTGT; 3' primer, CTCTATTGTTGGATCATAT. After the amplification, one fifth of the reaction mixture was slotted onto nylon filters and hybridized overnight with [³²P]-labelled oligomer probes in 5×SSPE, 5×Denhardt's and 0.5% SDS at 37 °C. The filter was then washed twice in 6×SSC at room temperature, followed by 20-min washes at 61 °C in 3 M tetramethylammonium chloride, 50 mM Tris pH 8.0, 2 mM EDTA pH 8.0, and 0.1% SDS. The final wash was in 5×SSC at room temperature.

DNA from the preleukaemic bone marrow of Ve showed tumours in two out of four attempts, with a latency of 43 days. When DNA from one of these primary nude-mouse tumours was used for secondary transfections, tumours formed in two out of two attempts, with a shortened latency period of 24 days. DNA from the patient's lymphoblastoid cell line gave no tumours in four transfections.

The nude-mouse tumours from initial transfections with bone marrow DNA from Ve carried human *Alu* repetitive sequences

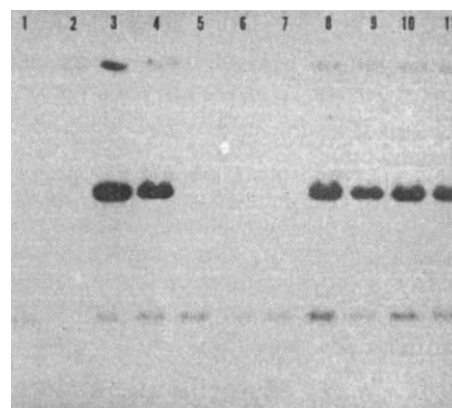


Fig. 3 A transforming allele of *Ki-ras* in patients with pre-leukaemia. DNA was extracted from nude-mouse tumours and then analysed by Southern blot hybridization for human *Ki-ras*. Samples: lane 1, control tumour bearing a mutant human *Ha-ras* gene; lane 2, control tumour bearing a mutant *N-ras* gene; lane 3, control tumour bearing a mutant *Ki-ras*; lane 4, tumour from transfection with DNA from patient BL; lanes 5–7, tumours without mutant *ras* genes; lanes 8–11, tumours from transfection with DNA from patient Mir.

(data not shown), and analysis with specific hybridization probes for oncogenes (*c-Ha-ras*, *N-ras* and *c-raf*) detected the presence of only one human oncogene: the Kirsten-*ras* oncogene *c-Ki-ras-2* (Fig. 1). The transforming capability of this *Ki-ras* allele was confirmed by demonstrating the presence of the gene in tumours formed by secondary transfections (Fig. 1).

Transforming *Ki-ras* alleles are usually mutated in codon 12 or 61 (ref. 6). We therefore cloned exons 1 and 2 of *Ki-ras2* from a secondary nude-mouse tumour and sequenced the pertinent portions^{7–9}. Unexpectedly, our results showed a glycine to aspartate mutation in codon 13, caused by a G to A transition in the second nucleotide of the codon (data not shown). Although the same base substitution has been found in the *N-ras* protooncogene from several acute leukaemias¹⁰, this is the first naturally occurring codon 13 mutation identified in *Ki-ras*.

We sought to determine whether the codon 13 mutation was present in the original bone marrow DNA of Ve. Using the polymerase chain reaction technique¹¹, we amplified the 100-base pair (bp) segment of the genomic *Ki-ras2* gene encompassing the 12th and the 13th codons in various DNA samples from Ve. These amplified segments were then probed with a 20-nucleotide oligomer specific for the codon-13 mutation. We found the codon-13 *Ki-ras* mutation in bone marrow and

Table 1 Activity of human DNA samples in a tumorigenicity assay

Source of DNA	Diagnosis*	Cytogenetics	Transforming gene	Conversion to acute leukaemia	Tumorigenicity†
Normal human tissues	—	—	—	—	1/13
T24 cell line	—	—	<i>Ha-ras</i>	—	7/7
Ve: bone marrow	RAEB	inv(3)(q21q26.2)	<i>Ki-ras</i>	Yes	2/4
Ve: lymphoblastoid cell line	—	Normal	—	—	0/4
Ve: primary nude-mouse tumour	—	—	<i>Ki-ras</i>	—	2/2
Mir: blood	RAEB	Normal	<i>Ki-ras</i>	Yes	1/4
BL: blood	ANLL from	inv(3)(q21q26.2), -7, i(17q)	<i>Ki-ras</i>	Yes‡	2/4
	RAEB				
So: bone marrow	RAEB	Normal	—	No	0/4
Lu: bone marrow	RA	Not done	—	No	0/4

* The diagnostic categories are in accordance with the French-American-British classification. RA, refractory anaemia; RAEB, refractory anaemia with excess blasts; RAEB in transition, refractory anaemia with excess blasts in transition; ANLL, acute nonlymphocytic leukaemia.

† Only those tumours that emerged within 60 days of the injection were considered positive.

‡ Blood sample used for analysis was obtained at the time the patient first presented with acute leukaemia.

peripheral blood cells collected from the patient at various times during the course of his disease, but not in DNA from his lymphoblastoid cells (Fig. 2). We conclude that the mutation in *Ki-ras* was present in original specimens from the patient and represented a somatic rather than germline lesion.

The possibility that other forms of genetic damage were involved in the establishment of this patient's preleukaemic state was examined by determining whether amplification or gross rearrangement of pertinent protooncogenes coexisted with the *ras* mutation. None of the genes tested (*abl*, *erb-B-2*, *Ha-ras*, *Ki-ras*, *myb*, *myc*, *N-ras* and *raf*) showed any abnormalities when examined by restriction mapping and Southern blotting (data not shown).

We applied the nude mouse assay to DNAs from the blood and bone marrow of three other patients (Mir, So and Lu) with MDS and one patient (BL) with acute leukaemia arising from MDS (Table 1). We found that two of these DNA samples (one MDS, one acute leukaemia from MDS) induced tumorigenicity. Analysis by Southern blotting revealed that the transforming gene in both patients was *Ki-ras* (Fig. 3 and Table 1). In total, two of four preleukaemics studied and one acute leukaemic whose disease arose from a preleukaemic state carried a transforming *Ki-ras* allele in their affected cells. It therefore appears that mutations in *Ki-ras* are present in the blood cells of certain patients with human preleukaemia. Moreover, all three of the patients in our series with *ras* mutations have progressed to frank leukaemia, whereas the two patients without detectable *ras* lesions have remained in a myelodysplastic state (Table 1).

The notion that mutant *ras* genes are involved in the initiation of tumorigenesis has been fostered by animal studies using chemical or viral carcinogenesis^{12,13} and transgenic mice¹⁴. Recently, mutations in *Ki-ras* have been uncovered in premalignant tissue adjacent to human colon carcinomas¹⁵, thus provid-

ing evidence for the involvement of mutant *ras* alleles in the initiation of a human malignancy. Our report, however, describes *ras* mutations in a human preleukaemic condition and documents the presence of the mutations in samples taken before the onset of frank malignancy. During the preparation of this manuscript, others¹⁶ have reported findings similar to ours, with the exception that their patients with MDS had mutations in *N-ras* rather than *Ki-ras*. Thus, the combined results from the two studies make it likely that genetic lesions in *ras* appear frequently in MDS and raise the possibility that the search for mutant *ras* genes may identify patients at high risk of conversion into acute leukaemia.

We thank Dr Marek Bozdech for bone marrow samples, Randy Saiki for the *Taq* polymerase, and Art Levinson for the plasmid pSVEKis. Supported by NIH grants, by funds from the G. W. Hooper Research Foundation; by an NCI Clinical Investigator Award and by the Rodney Foppiano Memorial Fund.

Received 17 June; accepted 17 August 1987.

- Greenberg, P. L. & Mara, B. *Am. J. Med.* **66**, 951-958 (1979).
- Bos, J. L. *et al. Blood* **69**, 1237-1241 (1987).
- Bitter, M. A. *et al. Blood* **66**, 1362-1370 (1985).
- Fasano, O. *et al. Molec. cell. Biol.* **4**, 1695-1705 (1984).
- Blair, D. G. *et al. Science* **218**, 1122-1124 (1982).
- Barbacid, M. *A. Rev. Biochem.* **56**, 779-828 (1987).
- McGrath, J. P. *et al. Nature* **304**, 501-506 (1983).
- Chen, E. Y. & Seeburg, P. H. *DNA* **4**, 165-170 (1985).
- Davis, L. G., Dibner, M. D. & Battey, J. F. *Basic Methods in Molecular Biology* 44-46 (Elsevier, New York, 1986).
- Bos, J. L. *et al. Nature* **315**, 726-730 (1985).
- Saiki, R. K. *et al. Science* **230**, 1350-1352 (1985).
- Zarbl, H., Saraswati, S., Arthur, A. V., Martin-Zanca, D. & Barbacid, M. *Nature* **315**, 382-386 (1985).
- Brown, K. *et al. Cell* **46**, 447-456 (1986).
- Quaife, C. J., Pinkert, C. A., Onniz, D. M., Palmiter, R. D. & Brinster, R. L. *Cell* **48**, 1023-1034 (1987).
- Bos, J. L. *et al. Nature* **327**, 293-297 (1987).
- Hirai, H. *et al. Nature* **327**, 430-432 (1987).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article
Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____
 Title _____
 Institution/Company _____
 Department _____
 Address _____
 City _____ State _____ Zip _____
 Phone (____) _____

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Liposome techniques in cell biology

C.R. Alving

The potential of liposomes as drug-delivery vehicles is currently receiving much attention. But new research uses are also being developed, in areas from modulating the phosphatidylinositol cycle to modelling antigen-presenting cell functions.

LIPOSOMES were first employed as carriers to introduce biologically active molecules into cultured cells in 1972. Since then, the field has proliferated rapidly, and a profusion of cell biological techniques using liposomes has emerged. Four types of liposome-cell interactions have been identified: stable adsorption (specific or non-specific), endocytosis, fusion, and lipid transfer¹.

Studies of the mechanisms of liposome-cell interactions have led to the development of many elegant applications for the delineation of subcellular architecture through the passage of liposome-associated probes. The first specific visualization of the Golgi apparatus in living cells is one of several notable achievements that were accomplished by using this technique².

Lipid metabolism

Modern cell biology is currently undergoing tremendous growth in the understanding of the functional roles of phospholipids and the metabolism of lipids. The phosphatidylinositol (PI) signal transduction cycle has been a central focus in many studies. A series of experiments by Wassef *et al.*^{3,4} demonstrated that PI metabolism in macrophages can be modulated by liposomes and liposome-associated PI. In these experiments, liposomes were first treated with a specific IgM antibody to a glycolipid antigen. Antibody binding to the liposomal glycolipid resulted in the coating of complement (C) on the surface of the liposomes. Liposomes prepared in this way are called liposome-antibody-complement (LAC). Upon exposure of macrophages to LAC, C-dependent phagocytosis occurred and this was accompanied by stimulation of the endogenous macrophage PI cycle (Fig. 1-I). Phagocytosis of LAC and stimulation of the macrophage PI cycle were both completely suppressed when the LAC liposomes themselves contained PI as one of the phospholipid ingredients in the liposomal bilayer (Fig. 1-II).

In the above system, when phosphatidylinositol-4-phosphate (PIP) was substituted for PI in the C-sensitized liposomes, no suppression of phagocytosis or endogenous macrophage PI metabolism was observed (Fig. 1-III). A monoclonal IgM antibody to PIP was manufactured and was used instead of the anti-glycolipid antibody to prepare LAC. PIP liposomes coated with anti-PIP antibodies and

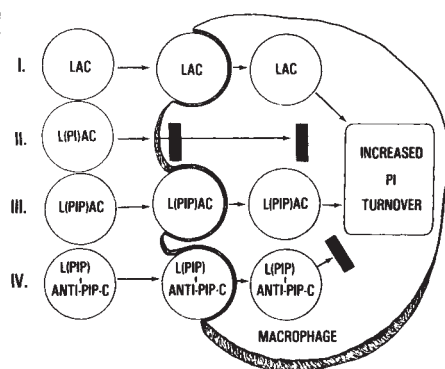


Fig. 1 Schematic illustration of the effects of liposomal PI, PIP, and antibodies to PIP on complement-dependent phagocytosis of liposomes and stimulated PI metabolism in macrophages.

complement underwent normal C-dependent phagocytosis, but this composition of LAC lacked the ability to induce increased turnover of PI in the endogenous macrophage PI cycle (Fig. 1-IV). It is evident that C-dependent phagocytosis and stimulation of the PI cycle are separate activities of macrophages that can be influenced by liposomes independently.

A unique method that uses liposomes containing phospholipase A₂ (PLA₂) substrates to kill tumour cells has been introduced⁵⁻⁷. An observation was made by Jett and Alving⁵ that liposomes containing plant PI exhibited potent killing activity against virtually all cultured malignant cell lines tested, but did not kill normal cells. Cytotoxicity against tumour cells was so great that the cells actually disintegrated and disappeared over a period of days. Liposomes containing PI derived from an animal source completely lacked tumoricidal activity.

Analysis of the metabolic fate of synthetic PI molecules in liposomes showed that those that had cytotoxic activity against tumour cells, *sn*-2 linoleoyl PI for example, were taken up by neuroblastoma cells as much as 10-fold more readily than synthetic PI molecules lacking cytotoxic activity, such as *sn*-2 oleoyl PI⁶. Nearly all of the fatty acids in the *sn*-2 position of plant PI consist of linoleic acid, while animal PI has little linoleic acid. The *sn*-2 linoleoyl fatty acid predisposes liposomal phospholipid molecules to PLA₂ degradation, and most of the degradation of labelled synthetic PI by transformed cells was accounted for by PLA₂. The products of PLA₂-induced degradation of PI are

lysoPI and free fatty acid, and when these latter molecules were presented to transformed cells they had the same devastating cytotoxic effects as liposomes containing plant PI. Although liposomes containing a variety of other phospholipids were not cytotoxic to tumour cells, the activity was not unique to PI. *Sn*-2 linoleoyl phosphatidylcholine also exhibited cytotoxic activity, albeit to a lesser degree⁷.

In this interesting tumoricidal system, the current theoretical killing mechanism suggests that malignant cells characteristically have a much higher inherent PLA₂ activity than nonmalignant cells: in fact, high PLA₂ activity may be a fundamental general property of malignant cells. Liposomes containing PLA₂ substrates therefore may be "targeted" to PLA₂ on tumour cells, and the resulting flood of local degradation products, including lyso compounds and free fatty acids, may induce rapid cytotoxicity.

The immune system

Since 1968 a considerable amount of research has been devoted to liposome models for studying immunologic phenomena. A currently evolving paradigm in immunology is that the afferent limb of the humoral immune response to many antigens involves processing of the antigen in specialized cells (often mononuclear phagocytes) referred to as antigen presenting cells (APC). After incorporation of antigen into the APC, the antigen becomes associated with class II (Ia) major histocompatibility complex (MHC) molecules. The combination of specific antigen and MHC is then transferred to the surface of the APC where it is presented for further processing by helper T lymphocytes that recognize the antigen-MHC combination. Walden *et al.*^{8,9} have now developed a model system in which liposomes carrying antigens can substitute for APCs to induce antigen-specific MHC-restricted interactions with lymphoid cells. The liposomes used in these model studies of humoral immunity, which might be termed "antigen-presenting liposomes" (APL), are the most recent products of an area of research that previously concentrated on development of APL in cell-mediated immunity for the stimulation of cytotoxic lymphocytes. This research complements work in other laboratories in which planar lipid membranes are used instead of liposomes.

In the APL system, the liposomes contained class II MHC molecules incorporated via their hydrophobic transmembrane region, together with a foreign antigen covalently linked to lipid. The appropriate liposomes in the absence of APCs were used to stimulate a proliferative response in an antigen-specific MHC-restricted T cell clone, or interleukin-2 production in an antigen-specific MHC-restricted T cell hybridoma. It is of interest that the foreign antigen and the MHC had to be in the same liposome to be effective, and the T cell response could be blocked by monoclonal antibodies directed against either the foreign antigen or the MHC molecules.

Another type of molecule that is critically important in certain aspects of the immunological functioning of the APC is a family of molecules called interleukin-1 (IL-1). IL-1 is secreted by monocytes or macrophages, and membrane-associated IL-1 may be important in the activation and proliferation of lymphocytes in the development of humoral immunity. IL-1-liposome complexes have now been constructed and tested for activity in an *in vitro* cell proliferation system¹⁰. IL-1 derived from monocyte plasma membranes was active in liposomes, but soluble IL-1 and IL-1 derived from the monocyte cytosol lost activity upon encapsulation in liposomes.

The above studies demonstrate that recent techniques using liposomes may have utility in many areas of cell biology. An example of a possible future employment of liposomes is in the induction of "anti-liposome" antibodies that react with cell membranes. Anti-liposome antibodies have specificities to particular liposomes, lipid bilayers, phospholipids, and phosphate esters, and under some conditions they might even cross-react with DNA¹¹. The ability of such antibodies to react with normally cryptic plasma membrane phospholipid antigens was recently demonstrated¹². □

Carl R. Alving is at the Department of Membrane Biochemistry, Walter Reed Army Institute of Research, Washington, DC 20307-5100, USA. For more information, fill in reader service number 100.

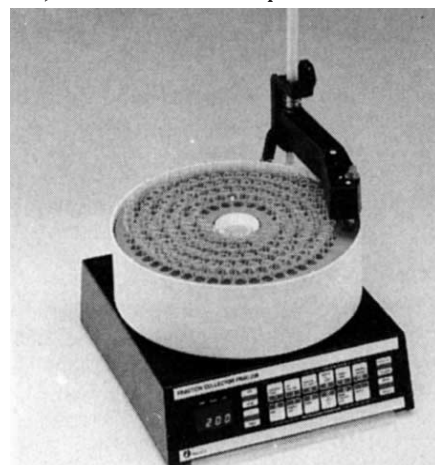
- Pagano, R.E. & Weinstein, J.N. *A. Rev. Biophys. Bioengng* **7**, 435-468 (1978).
- Lipsky, N.G. & Pagano, R.E. *Science* **228**, 745-747 (1985).
- Wassef, N.M., Roerdink, F., Richardson, E.C. & Alving, C.R. *Proc. natn Acad. Sci. U.S.A.* **81**, 2655-2659 (1984).
- Wassef, N.M. & Alving, C.R. *Meth. Enzym.* **141**, 244-255 (1987).
- Jett, M. & Alving, C.R. *Biochem. biophys. Res. Commun.* **114**, 863-871 (1983).
- Jett, M., Chudzik, J., Alving, C.R. & Stanacev, N.Z. *Cancer Res.* **45**, 4810-4815 (1985).
- Jett, M. & Alving, C.R. *Meth. Enzym.* **141**, 459-466 (1987).
- Walden, P., Nagy, Z.A. & Klein, J. *Nature* **315**, 327-329 (1985).
- Walden, P., Nagy, Z.A. & Klein, J. *Eur. J. Immun.* **16**, 717-720 (1986).
- Bakouche, O., Brown, D.C. & Lachman, L.B. *J. Immun.* **8**, 4256-4262 (1987).
- Alving, C.R. *Chem. Phys. Lipids* **40**, 303-314 (1986).
- Fogler, W.E., Swartz, G.M. & Alving, C.R. *Biochim. biophys. Acta* **903**, 265-272 (1987).

Cell biologists in St Louis

The 27th annual meeting of the American Society for Cell Biology will be held next week in St Louis, Missouri. Roughly 210 companies have booked exhibit booths, and some of the wares they will have on display are described below.

LEFT-SIDED entry **glove boxes** are new from Labconco, in booth 735 (*Reader Service No. 101*). Four models — models 50000, 50002, 50004 and 50350 — are equipped with the left-sided access door, which has an opening of over 370 square inches. All have 8-inch diameter glove ports, fitted with size 9¾ neoprene gloves that can be changed without contaminating the interior of the box. The viewing window of all models measures 16 × 31 inches and is sloped 30° to eliminate glare. Each box is equipped with two 3-wire, 115 V polarized and grounded receptacles, and an interior fluorescent light. The models 50000, 50002 and 50350 also have a magnehelic gauge to monitor negative pressure, and the models 50002, 50004 and 50350 come with an interchange chamber mounted on the left-side door. The prices of the glove boxes range from \$5,940 to \$7,330 (US).

Pharmacia has reserved booths 621-625, and will be featuring its PhastSystem for electrophoresis and its Protein A Superose chromatography columns for FPLC. Pharmacia says its new FRAC-200 **fraction collector** is useful in both FPLC and HPLC, as well as standard chromatography techniques (*Reader Service No. 102*). The FRAC-200 operates in time,



Pharmacia's FRAC-200 merry-go-round.

volume and drop counting collection modes, with either 95 or 175 fraction tube racks, or with microprep inserts for the collection of small volumes. Integration data can be obtained on up to 19 peaks per run with the FRAC-200, and can be read manually or transferred to a printer or a PC. The fraction collector can communicate with pumps, controllers for

automation, monitors and valves for peak cutting, and recorders for event marking.

In the near right-hand corner of the exhibit hall, in booths 1311 and 1313, Coulter will be showing off its Multisizer analyser with 25,600 channels of resolution for **counting and sizing cells** (*Reader Service No. 103*). New Accucomp software for the handling of particle size analysis data is now available for use with IBM PC AT/XTs and the £13,950 (UK) Multisizer. The software provides a wide variety of presentations, graphics, statistics and listings. Special routines for extrapolation, multi-tube analysis and data averaging are also included. Coulter says the new software is totally user-friendly, and requires no computer programming knowledge. Cables for linking the Multisizer with a computer are provided, plus a user's guide and instruction manual for program installation and data handling.

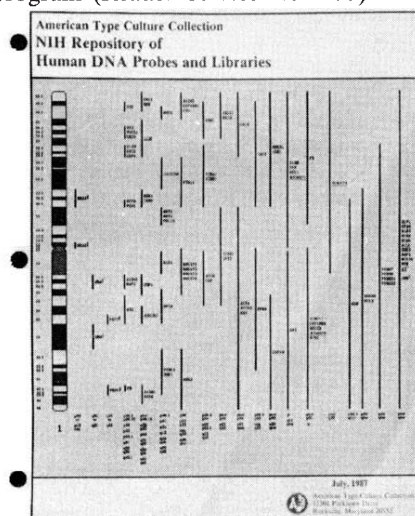
Ciba-Corning Diagnostics will have the Response II microprocessor-controlled **UV-visible spectrophotometer** on display in booth 324 (*Reader Service No. 104*). The £11,990 (UK) Response II is equipped with two 5.25-inch disk drives and comprehensive data manipulation facilities. No external interfacing with computers is required for operations such as comparison of samples and calculation of final results. Complex analysis procedures can be programmed into the instrument and stored for permanent use. Data generated during analyses can be transferred to or from disk and can be compared with previous results. The Response II can also be used with a rapid sampler accessory.

Sequencing and mapping

Beckman's exhibit will be hard to miss: it will take up two full islands directly across from the exhibit hall entrance, in booths 608 to 817. Beckman is giving star billing to its new GeneLine pulsed field electrophoresis system, for the separation of DNA fragments of up to nine million base pairs in length (*Reader Service No. 105*). The \$6,500 (US) system uses a new **transverse alternating field electrophoresis** technique where the electrodes are placed in front of and behind the gel, so that the DNA "zig-zags" through the thickness of the gel. Beckman says the GeneLine system produces perfectly straight lanes,

and can be used for separating and identifying whole chromosomes. One electrophoresis chamber is used for both electrophoretic separation and DNA blotting of fragile gels. The complete system includes the chamber with controller, a power supply, and a cooling system for controlling buffer temperature.

The American Type Culture Collection will have a display of its catalogues in booth 1040. One of the newest catalogues from ATCC is a 79-page description of **human DNA probes and libraries** deposited in the ATCC as part of the NIH Repository of Human DNA Probes and Libraries program (*Reader Service No. 106*). The



The book on human DNA probes and libraries. free catalogue contains probes and cloned genes sorted by chromosome and locus, gene name and locus, and depositor's name. Also included are human chromosome-specific genomic libraries, primate cDNA and genomic libraries, oncogene/transforming protein probes and clones sorted by locus name, and bacterial hosts for transformation or plating libraries. In addition, ATCC has listed related materials, such as certain oncogenes, that have been received through the ATCC general collection or the Patent Depository.

In the island across from Beckman, in booths 820-823, Du Pont will be extolling the virtues of its new **DNA sequencing system**, based on a modification of the Sanger chemistry (*Reader Service No. 107*). The Genesis 2000 system employs four fluorescent dyes developed from the family of succinylfluorescein compounds, covalently bound to chain terminators for the bases A, G, C and T. By coupling the dye to the terminator itself, Du Pont's technique requires only one reaction to tag all four nucleotides. Each dye produces different emission spectra when excited by the 488 nm line from an argon ion laser, that are picked up by two photomultiplier tubes. The ratio of the signals from each detector determines the

identity of the base. Du Pont says the system is more accurate than current systems, because prematurely terminated DNA fragments produce no signal. The system can run 12 lanes at once, with single-lane detection, and no special primers are needed for a throughput of 50-100 bases per lane per hour. The Genesis 2000 was introduced just three weeks ago, and Du Pont is now accepting orders for the \$90,000 (US) machine.

Four aisles to the left of the Du Pont exhibit, Boehringer Mannheim Biochemicals will be providing details on its **M13 sequencing kit** (*Reader Service No. 108*). Each kit contains all essential reagents, except radiolabelled nucleotides, for 100 sequencing reactions using the Sanger chain termination technique. 7-Deaza-dGTP is also included, to eliminate band compression of G-C rich regions of DNA. Boehringer Mannheim guarantees that a minimum of 250 nucleotides can be sequenced with the kit, and reports that typical results are in the 300-400 nucleotide range.

Breaking barriers

Also across the aisle from Beckman, Gibco/BRL will set up shop in booths 812-816. The Cell-Porator **electroporation system** will be the centrepiece of Gibco/BRL's instrumentation display (*Reader Service No. 109*). The \$1,199-1,399 (US) Cell-Porator system is designed for the transfection of mammalian or plant cells with DNA. The chamber safe holds four disposable electroporation chambers for



Porates cells with adjustable electric pulses.

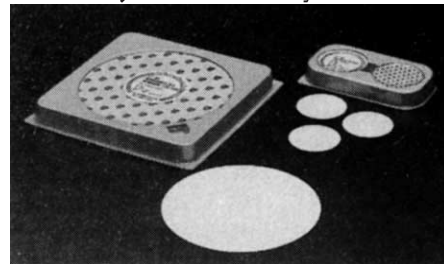
sequential electroporation. The pulse control unit is available with or without an internal power supply, and delivers a working voltage range of up to 400 V d.c., and reproducible pulses from 0.1 to 183 ms in length. Cells may be suspended in high or low salt buffers, and temperatures are kept constant by the insulated chamber safe. The system is supplied with 36 electroporation chambers and an instruction manual.

Lycozyme is useful for creating protoplasts from moulds, mushrooms and yeast. Accurate Chemical, in booths 620 and 622, sells **lycozyme** derived from *Trichoderma* (*Reader Service No. 110*). The b-1, 3-glucanase, chitinase and cellulase

enzymes in lycozyme degrade the cell walls of various groups of true fungi. Accurate Chemical sells lycozyme in 1 g quantities for \$95 (US).

Filter phenomena

Schleicher & Schuell has booked booths 638 and 640 in the middle rear of the exhibit hall. New from the company is a **large-pore cellulose acetate membrane filter**, with a mean pore size of 8 μm (*Reader Service No. 111*). Schleicher & Schuell says the filters may be used in



The S&S line-up of large-pore filters.

chemotactical assays, cytology procedures, parasitology studies and clinical diagnostics. The cellulose acetate composition of the membranes withstands treatment with alcoholic destaining solutions, making them useful in staining applications.

The Sterifil-D **disposable filter units** from Millipore are vacuum-operated devices for the sterilization of tissue culture media, microbiological media or other biological fluids in volumes of up to 150 ml (*Reader Service No. 112*). The units are available with low protein-binding, low-extractable polyvinylidene difluoride Durapore membranes, or surfactant-free nitrocellulose membranes. The Sterifil-D filters eliminate the need to transfer filtrate to storage containers: once the filter cup is removed, the receiver container can be stoppered for sterile storage. Millipore, exhibiting along the right-hand wall in booths 300-302, sells the Sterifil-D for \$36 (US) for 12 units.

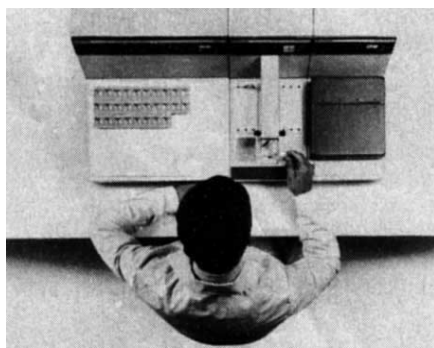
Nalge's booth, number 128, is situated near the entrance of the left-side restaurant. In its booth, Nalge will be displaying its newest filter item — Nalgene **disposable filter capsules** (*Reader Service No. 113*). Uses of the Nalgene filters include filtration and sterilization of tissue culture media, enzymes, buffers, and aqueous solutions. The filter capsules come with cellulose acetate and nylon membranes. The cellulose acetate filters feature a built-in prefilter and a double membrane with 0.8 and 0.45 μm , or 0.8 and 0.2 μm pore sizes. The cellulose acetate membranes exhibit low protein binding, and come in 500 cm² and 1,000 cm² sizes. Nalge says the nylon membranes are low in extractables, naturally hydrophilic, non-pyrogenic, and have no wetting agents. The nylon filters also come in 500 cm² and

1,000 cm² sizes, with pores of 0.2 and 0.45 µm. The filters range in prices from \$37.15 to \$52 (US), and operate under positive pressure.

Tissue specifics

Bio-Rad/Polaron's display in booths 528–533 will include the model E7480 **Cryo Freeze-drying System** for the preparation of samples for electron microscopy (*Reader Service No. 114*). The system removes water from specimens while preserving their essential structure. The unit has a diffusion-pumped vacuum system that functions at 10 mbar, and a specimen stage with a step-drying range from +20°C to –170°C. The model E7480 has a preparation chamber with the same time-saving advantages as the earlier model E7400, plus an airlock transfer mechanism. Sputter-coating and carbon-coating systems, and a film-thickness monitor are optional accessories.

Fisher Scientific offers a compact system for **paraffin-embedding tissues** for microscopy — the Histoprep embedding centre



Fisher's paraffin-embedding station.

(*Reader Service No. 115*). The \$4,600 (US) Histoprep consists of three individually controlled units: a thermal holding module, a dispensing module, and a cold-plate module. The modules are programmed to operate independently or as a unit. The centre is ergonomically designed so work flow can be switched from right-to-left or left-to-right, and the instrument has an event timer with 14 setpoints that includes a "standby function" to turn the centre on and prepare it for the day's work. Temperature sensors trip an audible alarm if temperatures vary by 1°C. The Histoprep has a 2-litre paraffin supply, storage for 350 moulds, and a cold-plate area that cools up to 50 moulds from ambient to selected temperature in 20 minutes. Fisher will be exhibiting in booths 208–216.

Growth enhancers

In the aisle opposite the left-side restaurant, in booth 225, Sigma Cell Culture will be introducing its plant tissue culture product line (*Reader Service No. 116*). The range of products includes 20 commonly used **plant tissue culture media** plus 40 plant cell culture-tested reagents and

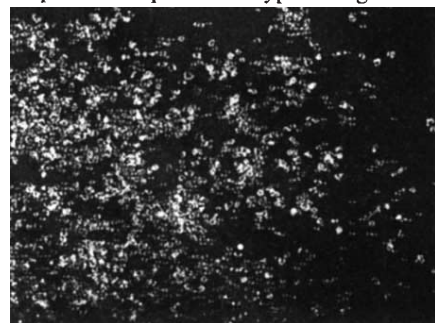


A range of plant media choices from Sigma.

biochemicals. Murashige and Skoog, Anderson's Rhododendron, Schenk and Hilderbrandt, and Nitsch and Nitsch are a sampling of some of the standard media available. Sigma also sells defined media prepared specifically for the micropropagation of particular plant species. Ten vitamin mixtures and solutions for the plant tissue culture scientist will be on display.

Mediatech Inc. will be showing its Cellgro line of liquid and powdered **cell culture media** in the near left-hand corner of the exhibit hall, in booth 108 (*Reader Service No. 117*). The Cellgro product line includes Ham's F-10 and F-12, Iscove's, McCoy's, DMEM, EMEM, and RPMI media, some supplied without glutamine. Salt solutions include Earle's, Hanks' and Dulbecco's, in 1× and 10× concentrations. Mediatech says its liquid media products are tested for the presence of mycoplasmas, and are supplied with endotoxin levels of less than 0.25 ng ml⁻¹.

In booth 1317 near the placement area, ICN Biomedicals will be displaying radiochemicals and biotechnology products. But ICN Biomedicals has also recently introduced a solid **support matrix for cell culturing**, called RapidCell (*Reader Service No. 118*). RapidCell is a specially-treated support matrix designed for substrate-dependent cells. ICN says RapidCell requires no trypsinizing for the



ICN Biomedicals' RapidCell support matrix for cell culture has lots of room to grow in.

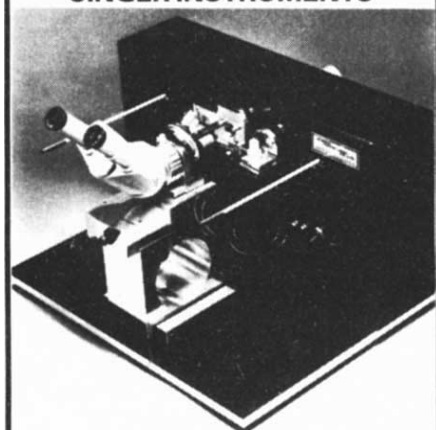
serial passage of cell lines, and that the controlled cell splitting saves flasks, culture dishes, pipettes, media and technician time. The higher cell densities and greater cell viability also increase the amount of excreted product per unit area and permit harvesting of more cells per volume, says the company. ICN Biomedicals sells the RapidCell matrix for \$20 (US) for 50 g.

VWR Scientific will have its **infrared CO₂ incubator** set up in booths 921 and 923 (*Reader Service No. 119*). The series 1800 automatic incubator has an infrared control system, capable of sensing a change in concentration as small as 0.1 per cent CO₂, says VWR Scientific. The incubator is available in either single-chamber or dual-chamber styles, for \$4,315 or \$6,985 (US), respectively. All models have an alarm to alert the user to changes in the interior CO₂ concentration. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further information, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

ADVERTISEMENTS

SINGER INSTRUMENTS



The Singer Microforge

For the production of: • MICRO-PIPETTES • NEEDLES
• HOOKS & LOOPS • ELECTRODE POLISHING

Tel: (0984) 40226

Telex: 337 492 Comcab G

Reader Service No.50

Versatile Tissue Culture Incubator

Modular Incubator Chamber (the Cell/Sphere™) for all tissue cultures, aerobic or anaerobic.

Low Cost



Eliminates evaporation loss and contamination. Easy to use — flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. billups-rothenberg, inc., pob 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.

Reader Service No.174